

Washington University in St. Louis Washington University Open Scholarship

All Theses and Dissertations (ETDs)

January 2011

Optimal Control of Inhomogeneous Ensembles

Justin Ruths

Washington University in St. Louis

Follow this and additional works at: <https://openscholarship.wustl.edu/etd>

Recommended Citation

Ruths, Justin, "Optimal Control of Inhomogeneous Ensembles" (2011). *All Theses and Dissertations (ETDs)*. 301.
<https://openscholarship.wustl.edu/etd/301>

This Dissertation is brought to you for free and open access by Washington University Open Scholarship. It has been accepted for inclusion in All Theses and Dissertations (ETDs) by an authorized administrator of Washington University Open Scholarship. For more information, please contact digital@wumail.wustl.edu.

WASHINGTON UNIVERSITY IN ST. LOUIS
School of Engineering and Applied Science
Department of Electrical and Systems Engineering

Dissertation Examination Committee:

Jr-Shin Li, Chair
Hiro Mukai
Joseph A. O'Sullivan
Heinz Schaettler
Sheng-Kwei Song
Victor Wickerhauser

OPTIMAL CONTROL OF INHOMOGENEOUS ENSEMBLES

by

Justin Arthur Ernest Ruths

A dissertation presented to the Graduate School of Arts & Sciences
of Washington University in partial fulfillment of the
requirements for the degree of

DOCTOR OF PHILOSOPHY

May 2011
Saint Louis, Missouri

ABSTRACT OF THE DISSERTATION

Optimal Control of Inhomogeneous Ensembles

by

Justin Arthur Ernest Ruths

Doctor of Philosophy in Systems Science and Mathematics

Washington University in St. Louis, 2011

Research Advisor: Professor Jr-Shin Li

This dissertation is concerned with formulating the problem and developing methods for the synthesis of optimal, open-loop inputs for large numbers of identically structured dynamical systems that exhibit variation in the values of characteristic parameters across the collection, or ensemble. Our goal is to steer the family of systems from an initial state (or pattern) to a desired state (or pattern) with the same common control while compensating for the inherent dispersion caused by the inhomogeneous parameter values. We compose an optimal ensemble control problem and develop a computational method based on pseudospectral approximations to solve these complex problems. This class of ensemble systems is strongly motivated by natural complications in the control of quantum phenomena, especially in magnetic resonance; however, similar structures are prevalent in a variety of other applications. From another perspective, the same methodology can be used to analyze systems that have uncertainty in the values of characteristic parameters, which are ubiquitous throughout science and engineering.

Acknowledgments

Let me first acknowledge and thank my advisor, Professor Jr-Shin Li. Throughout this work, he has consistently provided me with incredible projects and opportunities that have both challenged me and increased my curiosity and enthusiasm for research. I feel completely ready and excited about taking the next step in my career - and for that, which can never be taken away, I am sincerely grateful.

I would also like to thank my thesis committee members for their time and expertise as well as the rest of the ESE faculty. I feel privileged to have studied under such talented teachers and researchers. My lifelong academic goal will be to strive to match the level of excellence that each of you command. I also thank my collaborators, including Tsyr-Yan Yu for making my mathematical tinkering a physical reality and, most notably, Steffen Glaser for humoring my questions and freely sharing ideas.

I also thank all my fellow graduate students, in particular Juanyi Yu, Isuru Dasanayake, Anatoly Zlotnik, and Ji Qi as well as our postdoc Dionisis Stefanatos for their support, helpful discussions, and some occasional fun. Peter Lee has been a true friend to me since college, and I have enjoyed our shared, but separate graduate experience.

I am immeasurably fortunate to be where I am now and to have been given the encouragement and education to succeed. My parents, Mark and Rosemary, have provided unwavering support for all of my interests and are, at heart, responsible for the person I am today. My brothers, Derek, Troy, and Weston, have always been the source of inspiration and camaraderie that only brothers can supply. My wife, Melissa, shines so much joy into my life that it is easy to approach each day with eager anticipation. I can delve deeper and work harder knowing that she is there every step of the way.

Justin Arthur Ernest Ruths

Washington University in Saint Louis
May 2011

Dedicated to my family.

Contents

Abstract	ii
Acknowledgments	iii
List of Figures	vii
Preface	xi
1 Introduction	1
1.1 Inhomogeneous Ensembles	1
1.1.1 Quantum Ensembles	2
1.1.2 Biological Ensembles	4
1.1.3 Uncertain Systems	5
1.2 History of Pulse Design in NMR	5
1.3 Organization	6
1.4 Contribution	7
2 Ensemble & Optimal Control	9
2.1 Ensemble Controllability	10
2.1.1 Controllability of Bilinear Ensemble Systems	11
2.1.2 Controllability of Time-Varying Linear Ensemble Systems	15
2.2 Optimal Control	16
2.2.1 Solving Optimal Control Problems	17
2.3 Optimal Ensemble Control	20
3 Pseudospectral Method for Optimal Control	22
3.1 Spectral Methods	22
3.1.1 Comparison to Finite Element Methods	23
3.2 Pseudospectral Method	24
3.3 Example from Quantum Optics	28
3.4 Multidimensional Extension	29
3.5 Implementation	30
3.5.1 Visualization	31
3.5.2 Limitations	32
3.5.3 Programming Environments & Languages	32
4 Optimal Pulse Design for NMR & MRI	34

4.1	Pulse Design	34
4.2	Quantum Dynamics	36
4.3	Pulse Design on the Bloch Equations	37
4.3.1	Broadband Excitation and Inversion	38
4.3.2	Variation in Initial Conditions	44
4.3.3	Time-Varying Frequency	46
4.4	Relaxation Optimized Pulse Design	47
4.4.1	Polarization Transfer without Cross-Correlated Relaxation . .	48
4.4.2	Polarization Transfer with Cross-Correlated Relaxation	53
5	Convergence of the Pseudospectral Method	56
5.1	Empirical Convergence	59
5.2	Preliminaries	61
5.3	Main Result	65
5.4	Ensemble Extension	66
6	Conclusions	68
6.1	Future Work	69
Appendix A	Pseudospectral Method Supplement	71
A.1	Orthogonal Polynomials	71
A.2	Legendre Polynomial Properties for Optimal Control	72
A.3	Lagrange Interpolating Polynomials	73
A.4	Optimal Interpolation Nodes	73
A.5	Lagrange Polynomial written in terms of Legendre Polynomial	75
A.6	Derivative Matrix	76
Appendix B	AMPL Examples	78
B.1	Single Spin Bloch Optimization	78
B.2	Broadband Spin Bloch Optimization	80
Appendix C	Bloch Equations	83
C.1	Conversion	84
	References	86
	Vita	91

List of Figures

1.1	The effect inhomogeneity in the static magnetic field equally distributed around a center point (left) and in the electromagnetic pulse distributed equally around the nominal value (right).	4
2.1	The four-step back-and-forth maneuver of two vector fields, g_1 and g_2 , creates a displacement defined by, to the second order, the Lie bracket of the vector fields. If $[g_2, g_1]$ is not a linear combination of g_1 and g_2 , the Lie bracket synthesizes a new direction of control.	12
3.1	The time-optimal frictionless atom cooling problem in (3.10) is efficiently solved by the pseudospectral method. The dots in the two right-most figures indicates the 71 interpolation nodes used to approximate the original problem.	28
4.1	The control pulse shape (top), simulated resonance offset profile (middle), and experimental resonance offset profile (bottom) for an excitation (left) and inversion pulse (right) developed with the multidimensional pseudospectral method. The $\pi/2$ pulse achieves an average x component of 0.9852 with the parameters: $A = 20$ kHz, $B = 20$ kHz, $T = 50\mu\text{s}$, $N = 32$, $N_\omega = 8$. The π pulse achieves an average z component of -0.9991 with the parameters: $A = 20$ kHz, $B = 20$ kHz, $T = 120\mu\text{s}$, $N = 36$, $N_\omega = 12$. The difference between the simulated and experimental offset profiles highlights the effect of rf inhomogeneity (simulated response is shown for $\epsilon = 0.85, 0.9, 0.95, 1$).	40
4.2	A broadband π pulse and the corresponding excitation profile (red) compared to the excitation profile of the conventional hard pulse (black). The dips in the excitation profile are due to rf inhomogeneity in the experimental equipment. $A = 20$ kHz, $B = 40$ kHz, $T = 120\mu\text{s}$, $N = 36$, $N_\omega = 12$	40
4.3	The 2D spectra from on-resonance (black) and 250 ppm off-resonance (red) HSQC experiments using the pulse in Figure 4.2 (left), a standard adiabatic pulse (middle), and the hard pulse (right). The enhanced signal off-resonance enables better determination of protein structure and the optimized pulse achieves up to 20 times the sensitivity of the adiabatic pulse and with shorter duration ($120\mu\text{s}$ versus $500\mu\text{s}$). . . .	41

4.4	Arbitrary constraints are easy to include in the pseudospectral formulation for pulse design, such as with a symmetry constraint on the control or leaving the pulse duration free to vary, i.e. $0 < T \leq T_{\max}$. The average z magnetization of the offset (not shown) is -0.9923 . Parameters: $A = 20$ kHz, $B = 20$ kHz, $T_{\max} = 200\mu\text{s}$, $N = 24$, $N_{\omega} = 9$	42
4.5	Broadband excitation $\pi/2$ pulses designed to compensate for 10% rf inhomogeneity with average excitation 0.98. The minimum-energy broadband pulse (right) is optimized by solely minimizing energy subject to a performance constraint and achieves this transfer with 16% less rf energy. Parameters: $A = 20$ kHz, $B = 20$ kHz, $\delta = 0.1$, $T = 100\mu\text{s}$, $N = 24$, $N_{\omega} = 8$, $N_{\epsilon} = 1$	43
4.6	Broadband inversion π pulses designed to compensate for 10% rf inhomogeneity with average excitation -0.9929 . The pulse was optimized to minimize both magnetization and rf energy. Parameters: $A = 20$ kHz, $B = 20$ kHz, $\delta = 0.1$, $T = 120\mu\text{s}$, $N = 36$, $N_{\omega} = 8$, $N_{\epsilon} = 1$	43
4.7	Pulses are optimized to produce a desired $z \rightarrow y \rightarrow -y \rightarrow z$ evolution of the Bloch equations. The upper plot displays the concatenation of individually optimized $z \rightarrow y$ and $y \rightarrow -y$ pulses, which achieves the dashed terminal profiles shown below, with respective average performances: 0.99, 0.98, 0.97 (0.91 minimum). The middle plot displays a 3-part simultaneously-optimized pulse robust to variation in the initial condition and achieves the solid terminal profiles shown below, with respective average performances: 0.99, 0.99, 0.99 (0.97 minimum). The noticeable enhancement in performance and uniformity is due to compensating for the inhomogeneity in the initial condition of the individual pulses.	45
4.8	Control pulses (top) and state trajectories (bottom) corresponding to different objectives and designed to compensate for the time-varying frequency $\omega(t) = \sin(t)$. A single-system state transfer $M(0) = (0, 0, 1)' \rightarrow M(T) = (1, 0, 0)'$ is designed using the terminal cost $\varphi(T) = M_x(T)$ and running costs $\mathcal{L}(t) = 0$ (left), $\mathcal{L}(t) = 0.1(u(t)^2 + v(t)^2)$ (middle), $\mathcal{L}(t) = 0.1$ (right). The terminal time was free in all cases, bounded by $T_{\max} = 1$	46

4.9	(Left) The efficiency of the transfer $x_1 \rightarrow x_4$ in system (4.6) achieved by the pseudospectral method, as a function of the relaxation parameter ξ in the range $[0, 1]$. The theoretically calculated maximum efficiency given by (4.7) is also shown. (Right) ROPE and Pseudospectral controls (top) and corresponding state trajectories (bottom) for $\xi = 1$. Each of the hard pulses within the ROPE pulse at $t=0$ and $t=T$ (top left) correspond to a 35° rotation and transfer the state from $x(0^-) = [1, 0, 0, 0]^T$ to $x(0^+) = [\cos 35^\circ, \sin 35^\circ, 0, 0]^T$ and from $x(T^-) = [0, x_2(T), \eta \sin 35^\circ, \eta \cos 35^\circ]^T$ to $x(T^+) = [0, x_2(T), 0, \eta]^T$, respectively, in near instantaneous time.	50
4.10	The optimal ensemble pulses (right) effectively compensate for all variations of ξ on the interval $[0, 2]$ with only minor losses in transfer efficiency (left) when compared to each analytic ROPE [32] efficiency for a single value of ξ ($N = 28$ and $N_\xi = 8$). The top optimal ensemble pulse was developed by maximizing the average transfer efficiency and the pulse beneath was developed by maximizing the average transfer efficiency and minimizing energy. The transfer efficiency plot corresponds to the latter.	52
4.11	The optimal ensemble pulse shown in (right) effectively compensates for all variations of $\xi \in [0, 2]$ and $J \in [0.5, 1.5]$ with comparable efficiency (left) to each ROPE pulse [32] for a specific ξ and J . This optimal ensemble pulse was developed by maximizing average transfer efficiency and minimizing energy with $N = 24$, $N_\xi = 8$, and $N_J = 4$. .	52
4.12	The efficiency (left) of the transfer $x_1 \rightarrow x_6$ in system (4.9) without variation achieved by the pseudospectral method, as a function of the relaxation parameter ξ_a in the range $[0, 1]$, with $\xi_c = 0.75\xi_a$. The efficiency (right) of the ensemble case closely reproduces the optimal efficiency with a single pulse for the entire $[0, 1]$ range.	54
5.1	Numerical results are shown for the convergence of the pseudospectral transfer efficiency to the optimal transfer efficiency in the cases of the polarization transfer problems (4.8) for $\xi = 1$ (left), and (4.9) for $\xi_a = 1$ and $\xi_c = 0.75$ (right). The error in these examples is the difference between the pseudospectral transfer efficiency and the analytic optimal efficiency.	60
5.2	The numerical convergence of the Bloch system is shown in both the time and parameter domains. The performance of the optimal broadband $\pi/2$ pulse converges to unity as the number of discretizations, N , or samples, N_ω , independently get large. (Parameter values: maximum rf amplitude = 20 kHz; bandwidth = $[-20, 20]$ kHz; duration = $100\mu s$; fixed $N_\omega = 8$, left; fixed $N = 30$, right).	60

A.1	Interpolation approximations of the function $f(t) = 1/(16t^2 + 1)$ using an LGL and uniform grid, respectively, with $N = 16$	73
A.2	The dramatic difference in interpolation error apparent in the Lebesgue constant for interpolations based on LGL and uniform grids.	74

Preface

The escalation in both our understanding of the natural world and our technology to construct engineered mechanisms has led to increasingly more complex models of important dynamical systems. This dissertation attempts to bridge the widening gap between this complexity and our capacity for mathematical analysis by developing new tools that provide a systematic approach to characterize and control special classes of complex dynamical systems. This type of work is a relatively rare art and often left out of multidisciplinary teams. However, it is research like this that is the glue between engineering, mathematics, and science - fully integrating the competencies of the separate efforts. Moreover, in this age of efficiency it is of even greater importance to fully understand the limits of our current models and methods so that we can squeeze the last drop out of every system that we create, engineer, and study.

It is with this awareness and perspective that the value of this work and others like it should be understood. With this in mind, my dissertation but scratches at the surface of the grand challenge facing future innovation and research. However, the stepwise methodology as well as the close integration of theory and computation is a strong model to sustainably keep pace with the exciting advancements in science and engineering.

Chapter 1

Introduction

In this work, we will develop a new framework and novel methods to synthesize optimal, open-loop controls for a class of complex dynamical systems called inhomogeneous ensembles. Such systems are motivated by the challenges arising from pulse design problems in nuclear magnetic resonance (NMR) and imaging (MRI), but also appear in additional areas of quantum control as well as science and engineering. Studying the conditions for the controllability and developing methods for computing optimal control laws of optimal ensemble control problems will foster many new avenues of research in theoretical control, systems theory, and computational methods. In the following sections I discuss in more detail the motivation and background for this research area.

1.1 Inhomogeneous Ensembles

The boundaries of scientific discovery and engineering interest lie at the brink of human - and mathematical - comprehension. Attempting to accurately model the millions upon millions of interacting cells, the complex network of protein pathways at the intra-cellular level, or the Avogadro's number of quantum spins involved in typical experiments yields models of tremendous scale. In many of these applications feedback is either extremely limited, impractical, or expensive due, in part, to the sheer number of the systems involved. Our desire to engineer and control leads us to develop a manageable mathematical framework to analyze such large-scale complex systems. This thesis is concerned with creating a systematic approach towards optimal control problems of a special class of complex systems, in which the dynamics of

the many sub-members can be indexed as a function of specific system parameters. In the next section I describe the original motivation for inhomogeneous ensembles from quantum systems. Since its introduction, however, new promising applications have been identified in other key areas of research and suggest a new and expanding topic in systems analysis and control theory.

1.1.1 Quantum Ensembles

Quantum science represents a diverse frontier of research and promises to deliver substantial advancements in, but not limited to, early-stage cancer detection, protein structure and function, drug delivery, as well as quantum information, computation, and optics [67, 12, 44, 24]. Common amongst these varied applications is the small size (nanometers) and short timescales (microseconds) involved, which require researchers to use elaborate equipment to interact with these ultra-small and ultra-fast systems. In all quantum experiments, electromagnetic signals are used to manipulate the quantum elements by shaping the system Hamiltonian (see Section 4.2 for more details). Designing these signals, or pulses, intelligently can lead to significant improvements in signal recovery, which in turn results in increased performance, efficiency, or resolution depending upon the application. From the perspective of control theory, this design goal can be cast as a state transfer problem of a bilinear (linear in both state and control) dynamical system. The synthesis of these controls is typically an open-loop process, as measurement of the system state for feedback is generally either unavailable or expensive to obtain.

At the quantum level, there are no direct interactions; instead, the behavior is dictated entirely by the interplay of the electromagnetic fields surrounding the quantum systems. In practice, these field strengths are highly susceptible to various forms of inhomogeneity originating from equipment irregularity as well as chemical interactions [15]. Compensating for the dispersive effect of these inhomogeneities in the pulse design process is key to preserving the signal recovered in a realistic experiment and is the motivation for the mathematical and computation framework developed here.

A canonical quantum spin system from NMR is given by the Bloch equations, in which the time evolution of the spin magnetization vector M is modeled by

$$\frac{d}{dt} \begin{bmatrix} M_x \\ M_y \\ M_z \end{bmatrix} = \begin{bmatrix} 0 & -\omega & u \\ \omega & 0 & -v \\ -u & v & 0 \end{bmatrix} \begin{bmatrix} M_x \\ M_y \\ M_z \end{bmatrix},$$

(see 4.3 for complete formulation). The behavior is characterized by a rotation around the effective magnetic field, which is a combination of the static magnetic field applied in the $+z$ direction and the electromagnetic pulse applied in the transverse plane. The targeted spin is typically a particular atom (e.g., hydrogen in proton NMR or in MRI), which has a specific characteristic physical constant called the gyromagnetic ratio (γ). The angular velocity, ω_{eff} , about this effective field, \mathcal{B}_{eff} , is given by the product $\omega_{\text{eff}} = \gamma \mathcal{B}_{\text{eff}} = (v, u, \omega)'$. Although the gyromagnetic ratio is unique and fixed for a given atom, the magnetic field is susceptible to variation, i.e., $\mathcal{B}_{\text{eff}} \in [\mathcal{B}_{\text{nom}} - \Delta, \mathcal{B}_{\text{nom}} + \Delta]$ is distributed about a nominal value. The major sources of this variation stem from equipment irregularities, which, for example, cause strong fields nearer to the source and weaker fields farther from the source. In addition, surrounding molecules can attenuate the effective magnetic field through chemical shielding. These so-called chemical shifts are crucial tools in NMR spectroscopy in determining the chemical structure of complex molecules. In some applications, however, and also in MRI, in which hydrogen atoms near regions of fat or bone are shifted differently, there is great need to compensate for this chemical shielding as well.

The difference in the trajectories taken by atoms experiencing slightly different effective magnetic fields accumulates and can cause very noticeable dispersion in the final states of the spins. Figure 1.1 shows the dramatic difference in the state trajectories, starting at $(0, 0, 1)'$ attempting to reach $(1, 0, 0)'$ under the control of an on-resonance hard pulse ($u(t) = A$ and $v(t) = 0$, where A is a constant amplitude), due to variation in the static field and applied pulse, respectively. Compensating for these effects then means developing pulses that will steer all of these inhomogeneous spins between these points of interest simultaneously. Because there is no way to apply a separate pulse to each of the spins, this pulse is common to all members of the ensemble.

From a systems and control perspective, this poses a very interesting and challenging problem where we model these inhomogeneous effects as a family of dynamical systems

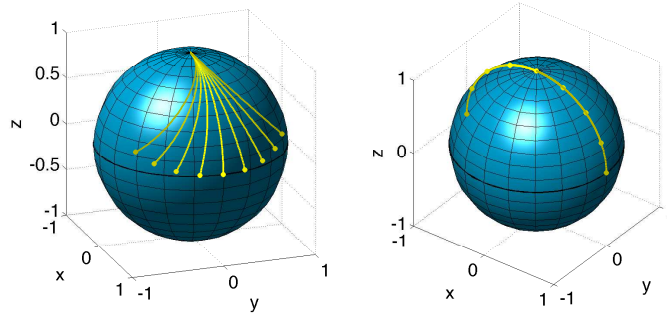


Figure 1.1: The effect inhomogeneity in the static magnetic field equally distributed around a center point (left) and in the electromagnetic pulse distributed equally around the nominal value (right).

with variation in the value of characteristic parameters. Designing compensating pulses is then a highly under actuated open-loop steering problem on an infinite dimensional state space. This forms a new area of systems theory directly related to quantum control, but with wider applications to other areas within science and engineering.

1.1.2 Biological Ensembles

The large volume of cells within living tissue makes modeling their behavior a fundamental challenge to studying and optimizing these systems. In such cases, there are millions of cells that operate in roughly the same way, however, not entirely identical. For example, the dynamics of an action potential of neuronal cells follows a cyclic process, which can receive input from surrounding cells or stimuli. It is possible to extract a fundamental baseline (zero-input) frequency of oscillation by linearizing about this high-dimensional limit cycle [47]. The characteristic frequency of oscillation of “firing” tends to vary across the population of neurons, so it is well modeled as an inhomogeneous ensemble. The synchrony and de-synchrony of these neurons is directly linked to conditions like Parkinson’s disease and epilepsy. Therefore, studying and controlling the ensemble synchronization patterns is of keen clinical interest. Such structures of large numbers of varied entities is repeated in many parts of biology, including cardiac tissue and protein signaling pathways within the cell.

1.1.3 Uncertain Systems

In the systems mentioned above there is a physical population of parallel structures that show a measurable variation, e.g., different frequencies. However, the same methodology can be used to model and control single systems that evidence uncertainty in one or more of the parameters within the dynamics. In an analog to the preceding example, an engineered system might have uncertainty in the stiffness of a beam or strut. This can be modeled as a spring with uncertainty in the spring constant k (i.e. $\ddot{x} = -kx$). Although there is actually only a single value that corresponds to the true spring constant, using the methods for inhomogeneous ensembles, we can study and derive controls to drive the system which will work for all values of spring constant within the specified uncertainty. This technique will be able to provide improvements in, for example, manufacturing robotics, flight control, and sensorless robust navigation. Recently such analysis has yielded a single ensemble control that steers a collection of nonholonomic vehicles with different wheel radius and that has been verified experimentally [1].

1.2 History of Pulse Design in NMR

While there is little precedent for control of mathematical ensembles, those working in pulse design have unwittingly approached this problem for many years. Pulse design has been and continues to be an active area of research with a rich and diverse literature on the topic. A majority of this work, especially done in the early history of NMR, is based on intuition and others on ad-hoc methods to compensate for the inherent inhomogeneity. Recently, interest in optimal control has provided a new and reliable tactic for tackling these problems. The emphasis of this dissertation is to additionally provide a clear mathematical framework for posing pulse design problems that includes variation in parameter values.

Quickly after the discovery of magnetic resonance and chemical shift effects, a variety of intuitive methods were developed and a suite of elementary pulses became common in NMR experiments. One of the techniques, spin echoes, uses a reverse evolution of a bandwidth of spins to undo the dephasing caused by the forward evolution [25, 11].

More elaborate designs then used combinations of these fundamental pulses to yield composite pulse sequences, which again took advantage of the inherent symmetry involved in the evolution of the spins and the Bloch sphere [35, 59, 66]. Today there exist large volumes of these pulses that have been assembled in clever ways to try to mitigate the effect of the inhomogeneities experienced in a real experiment [2]. However, it was soon realized that intuition was not sufficient to maximize the performance in quantum systems and ad-hoc, iterative algorithms were employed to design pulses [60]. Soon thereafter the analytic Shinnar-Le Roux selective excitation algorithm was developed based on the idea of small tip-angle approximations [50]. Although optimal control has origins in pulse design back to the 1980s and 1990s [13, 53], it has been a major contributor to pulse design methods in the past decade [61, 28, 33]. A family of computational methods have been applied to these problems including the gradient and Krotov methods [31, 46].

This dissertation offers a unique perspective on the design of pulses for NMR applications by providing a unified method from start (modeling and analysis) to finish (computation and synthesis). In particular, the mathematical characterization of the problem suggests new approaches for the numerical methods, which we follow in this work.

1.3 Organization

This dissertation is organized as follows:

- In Chapter 2, I provide a more rigorous definition of ensemble control and existing results on ensemble controllability. I review key aspects of optimal control that will be used throughout this dissertation and then formulate an optimal ensemble control problem.
- In Chapter 3, I introduce a pseudospectral method developed to reduce optimal control problems to nonlinear programming problems. I then show a direct extension of the method that accommodates optimal ensemble control problems. I also touch upon several practical topics including implementation and limitations.

- In Chapter 4, I present several examples of optimal ensemble control problems from quantum pulse design and the corresponding solutions found by the multidimensional pseudospectral method. I consider a unitary quantum system modeled by the Bloch equations and solve several practical examples including inhomogeneity in static and applied fields, variation in initial conditions, and time-varying frequency. I then focus on two versions of a non-unitary quantum system with and without the consideration of ensemble effects. Analytic optimal solutions are known for these problems and they provide a benchmark to use when evaluating the controls found by the pseudospectral method.
- In Chapter 5, I address the convergence of the single- and multidimensional pseudospectral method for optimal control and optimal ensemble control problems, respectively. In particular, I show that the solutions of the nonlinear programming problem generated by the pseudospectral method converge to the optimal solutions of the original problem as the order of approximation (discretization) increases. We provide this proof for the pseudospectral method and then provide extensions to consider the ensemble case. Several empirical results from problems in Chapter 4 are shown to motivate the analysis.

1.4 Contribution

The novelty of this dissertation belongs to the extensions to consider the optimal ensemble control. In particular,

- **Theory:** This thesis contributes an all-new formulation of optimal control problems for ensemble systems in Section 2.3. Chapter 5 presents a new approach to prove the convergence of the pseudospectral method for optimal control as well as the extension to the multidimensional pseudospectral method for optimal ensemble control.
- **Computation:** The extension of the pseudospectral method to consider the optimal ensemble control problem in a systematic and rigorous fashion is the central computational contribution of this work. In addition, the implementation choices to streamline the use of the pseudospectral and multidimensional

extension, such as using AMPL, is not preceded in the literature and will allow the method to reach a broader audience by being able to incorporate it into an online webservice.

- **Application:** Prior to this work, the application of the pseudospectral method was largely limited to trajectory-type problems motivated by, e.g., satellites. Applying, and extending, this methodology to the control of quantum systems is a unique to this thesis. A majority of work in optimal control of quantum systems relied on variations of the standard gradient method. All of the results presented in Chapter 4 are novel to this thesis work, as well as the example presented in Chapter 3.

Chapter 2

Ensemble & Optimal Control

The types of systems discussed in the preceding chapter motivate us to devise a systematic way to characterize and study inhomogeneous ensembles. Moreover, the practical constraints imposed by safety concerns and hardware limitations highlight the importance of being able to solve problems based on these ensemble systems in an optimal manner. In this chapter, I present an overview of the theory of ensembles, which guides the development of the rest of this work. I also review standard results in optimal control and present the formulation for optimal control of ensemble systems.

Although the introduction of ensemble control has been quite recent, related work has been studied for a number of years. The challenges inherent in the control of quantum systems, as presented in the Introduction, has been a longstanding problem in chemical physics and has yielded many interesting problems in systems and control theory. Some of these applications were first interpreted from the perspective of stochastic control [5, 8]. In these cases the sub-members of the ensemble were characterized by identical dynamics, however, influenced by different Brownian motions. Subsequently, these problems were then cast as a new form of control of parameterized control systems, which required a novel controllability analysis. The inhomogeneous Bloch equations were the prototype system of ensemble control, followed by other examples from the control of quantum systems [38, 36]. Ensemble controllability was then derived for time-varying, finite-dimensional linear systems using a different methodology. Below we expand upon these concepts, which form the foundation for the presented work in optimal ensemble control.

2.1 Ensemble Controllability

Ensemble control [37] is a mathematical framework to characterize parameterized dynamical systems of the form,

$$\frac{d}{dt}x(t, s) = F(t, s, x(t, s), u(t)), \quad x(0, s) = x_0(s), \quad (2.1)$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $s \in \mathcal{S} \subset \mathbb{R}^d$, with F and $x_0(s)$ smooth functions of their respective arguments. The significant challenge of this class of control problems originates from requiring the same open-loop control, $u(t)$, to guide the inhomogeneous continuum of systems from an initial distribution, $x_0(s)$, to a desired final distribution over the corresponding function space. We require such a framework to analyze and control ensemble systems in a systematic fashion. As can be seen in the quantum pulse design literature, there are dozens of ad hoc approaches to these problems that attempt to use intuition, however, this type of work is more guess-and-check than rigorous engineering. It is only in the context of a mathematical framework that a dependable metric for optimality can be established and bounds on performance and efficiency can be understood.

We use ensemble control to investigate the pertinent fundamental properties of ensemble systems. The first natural question that arises is whether such systems are controllable at all. The aim is to compensate for the dispersion inherent in the parameterized dynamics; however, there is no guarantee that it is possible. Therefore, by taking the form of the dynamics and the parameters into account, ensemble controllability assesses the types of variations that can be compensated [36].

Definition 1. *The family of systems in (2.1) is called **ensemble controllable** on the function space $\mathcal{F}(\mathcal{S})$ defined on some compact set $\mathcal{S} \subset \mathbb{R}^d$, if there exists a control law $u(t)$ such that starting from any initial state $x(0, s)$ the system can be steered to within a ball of radius ϵ around the target state $g(s) \in \mathcal{F}(\mathcal{S})$, i.e., $\|x(T, s) - g(s)\| < \epsilon$. Here $\|\cdot\|$ denotes a desired norm, say L_2 norm, on $\mathcal{F}(\mathcal{S})$. The final time T may depend on ϵ .*

As a new field there is much left wide open for investigation. The pioneering work in this area investigated the controllability of linear time-varying ensemble systems and

the bilinear Bloch equations with inhomogeneity, which as mentioned earlier was the prototype for this field of study [37, 38]. Results for general nonlinear systems are of keen interest as well as the quantification of reachable sets, which defines the possible controllable states in the absence of complete controllability. Moreover, advancements in the analysis of these systems will promote new methods, as in this work, for control and optimal control of ensembles.

2.1.1 Controllability of Bilinear Ensemble Systems

In what follows we consider the control affine dynamical system,

$$\frac{d}{dt}x = f(x) + \sum_{i=1}^m u_i g_i(x), \quad (2.2)$$

where f and g_i are the drift and control vector fields that drive the motion of the state, x , on a manifold $M \subset \mathbb{R}^n$ with controls u_i , for $i = 0, 1, \dots, m$. The study of the controllability of such systems can be understood using the tools of Lie groups.

A fundamental mathematical concept in this area is the Lie bracket, which represents possible directions of control that might not be in the linear span of the vector fields. It can be understood as follows. Consider the driftless system

$$\frac{d}{dt}x = g_1(x)u_1 + g_2(x)u_2,$$

where u_1 and u_2 are the controls which can be used to steer the system. Given a small unit of time t , we can propagate the system forward for the sequence of four control steps, each with length t : $(u_1, u_2) = (1, 0) \rightarrow (0, 1) \rightarrow (-1, 0) \rightarrow (0, -1)$. The evolution yields $x(4t) = (\exp -tg_2)(\exp -tg_1)(\exp tg_2)(\exp tg_1)x(0)$. If we expand this expression, to the second order we find,

$$x(4t) = x(0) + t^2[g_2, g_1] + O(t^3),$$

where

$$[g_2, g_1] = \frac{\partial g_2}{\partial x}g_1 - \frac{\partial g_1}{\partial x}g_2,$$

is the Lie bracket of g_1 and g_2 [36]. This quantity then represents, to the second order, the difference in initial and final positions $x(4t) - x(0)$. If $[g_2, g_1] \notin \text{span}\{g_1, g_2\}$, then this back-and-forth maneuver generates a new direction of controllable motion. Supposing this is true, we can continue these Lie bracket calculations between the now three vector fields available $\{g_1, g_2, [g_2, g_1]\}$. Therefore, the study of controllability becomes whether enough independent new directions can be synthesized through the appropriate nested Lie brackets of the original vector fields. This framework was developed for both the driftless (as above) and with-drift systems [7, 27].

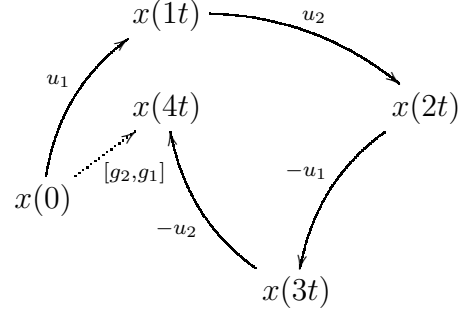


Figure 2.1: The four-step back-and-forth maneuver of two vector fields, g_1 and g_2 , creates a displacement defined by, to the second order, the Lie bracket of the vector fields. If $[g_2, g_1]$ is not a linear combination of g_1 and g_2 , the Lie bracket synthesizes a new direction of control.

Finite Number of Bloch Systems

The framework of Lie brackets allows us to approach the problem of a finite number of Bloch systems controlled by the same open-loop input. Specifically, consider the inhomogeneous Bloch equations,

$$\frac{d}{dt}M(t, \omega, \epsilon) = \left[\omega \Omega_z + \epsilon u(t) \Omega_y + \epsilon v(t) \Omega_x \right] M(t, \omega, \epsilon),$$

where $\omega \in [-B, B]$ and $\epsilon \in [1 - \delta, 1 + \delta]$, $0 < \delta < 1$, are parameters exhibiting variation (see 4.3 for complete description and Appendix C for the definitions of the generators of rotation Ω_α). For now suppose we sample these parameters resulting in the finite collection of Bloch systems,

$$\frac{d}{dt}M_{ij}(t) = [\omega_i \Omega_z + \epsilon_j u(t) \Omega_y + \epsilon_j v(t) \Omega_x] M_{ij}(t), \quad (2.3)$$

where ω_i distinct, $i = 1, 2, \dots, N_\omega$, are sampled from $[-B, B]$, and ϵ_j distinct, $j = 1, 2, \dots, N_\epsilon$, are sampled from $[1 - \delta, 1 + \delta]$. Combining this set of samples into a large combined system, it forms a $3N_\omega N_\epsilon$ -dimensional control system evolving on a Lie group $G \equiv (\text{SO}(3))^{N_\omega N_\epsilon}$ by

$$\dot{\Theta} = H\Theta, \quad (2.4)$$

where $\text{SO}(3)$ is the special orthogonal group and $\Theta \in G$; $H = \bigoplus_{i,j} H_{ij}$, the direct sum of 3×3 matrices $H_{ij} = \omega_i \Omega_z + u \epsilon_j \Omega_y + v \epsilon_j \Omega_x$, $i = 1, 2, \dots, N_\omega$, $j = 1, 2, \dots, N_\epsilon$, and $H \in \mathfrak{g}$, the Lie algebra of G . In this finite case, since G is compact and connected, if the Lie algebra generated by the vector fields $\{\omega_i \Omega_z, \epsilon_j \Omega_y, \epsilon_j \Omega_x\}_{LA} = \mathfrak{g}$, then every point in the group G can be reached by certain control input, which is analogous to the vector fields spanning the entire space as detailed in the previous section. We can express the recursive Lie bracket as,

$$\text{ad}_X^k(Y) = [X, \text{ad}_X^{k-1}(Y)]$$

for any $k \geq 1$, with $\text{ad}_X^0(Y) = Y$. Performing recursive Lie brackets on the sampled system in (2.3) leads to

$$\begin{aligned} \text{ad}_{\epsilon_j \Omega_y}^{2\ell}(\text{ad}_{\omega_i \Omega_z}^{2k}(\epsilon_j \Omega_x)) &= (-1)^{k+\ell} \omega_i^{2k} \epsilon_j^{2\ell+1} \Omega_x, \\ \text{ad}_{\epsilon_j \Omega_y}^{2\ell+1}(\text{ad}_{\omega_i \Omega_z}^{2k}(\epsilon_j \Omega_x)) &= (-1)^{k+\ell+1} \omega_i^{2k} \epsilon_j^{2\ell+2} \Omega_z, \\ \text{ad}_{\epsilon_j \Omega_x}^{2\ell}(\text{ad}_{\omega_i \Omega_z}^{2k}(\epsilon_j \Omega_y)) &= (-1)^{k+\ell} \omega_i^{2k} \epsilon_j^{2\ell+1} \Omega_y, \end{aligned}$$

where $k, \ell \in \mathbb{Z}_{\geq 0}$ are nonnegative integers. Now, let

$$\begin{aligned} X_{k\ell} &= \{\omega_i^{2k} \epsilon_j^{2\ell+1} \Omega_x\}, \\ Y_{k\ell} &= \{\omega_i^{2k} \epsilon_j^{2\ell+1} \Omega_y\}, \\ Z_{k\ell} &= \{\omega_i^{2k} \epsilon_j^{2\ell+2} \Omega_z\}, \end{aligned}$$

where $i = 1, 2, \dots, N_\omega$, $j = 1, 2, \dots, N_\epsilon$ and $X_{k\ell}, Y_{k\ell}, Z_{k\ell} \in \mathfrak{g}$. For $k = 1, 2, \dots, N_\omega$ and $\ell = 1, 2, \dots, N_\epsilon$, it can be shown that $\{X_{k\ell}, Y_{k\ell}, Z_{k\ell}\}$ form a linearly independent set if $\omega_m \neq \omega_n$, $m \neq n$, and $\epsilon_p \neq \epsilon_q$, $p \neq q$, for $m, n = 1, 2, \dots, N_\omega$ and $p, q = 1, 2, \dots, N_\epsilon$ [36]. Therefore,

$$\text{span}\{X_{k\ell}, Y_{k\ell}, Z_{k\ell}\} = \mathfrak{g},$$

and the finite family of $N_\omega N_\epsilon$ Bloch systems in (2.3) is controllable.

Ensemble Bloch Equations

The fundamental challenge for the controllability of ensembles is that we represent the family of systems as an uncountable continuum. For this reason the preceding methodology does not apply because the ensemble system evolves on an infinite dimensional Lie group. The approach for the ensemble case is most transparent for the driftless Bloch equations, i.e., no variation in the frequency parameter ω ,

$$\frac{d}{dt}M(t, \epsilon) = \epsilon \left[u(t)\Omega_y + v(t)\Omega_x \right] M(t, \epsilon).$$

Similar to above we can investigate the Lie algebra generated by the control vector fields $\{\epsilon\Omega_y, \epsilon\Omega_x\}$,

$$\begin{aligned} \text{ad}_{\epsilon\Omega_y}^{2k-1}(\epsilon\Omega_x) &= (-1)^k \epsilon^{2k} \Omega_z, \\ \text{ad}_{\epsilon\Omega_y}^{2k}(\epsilon\Omega_x) &= (-1)^k \epsilon^{2k+1} \Omega_x, \end{aligned}$$

where $k \in \mathbb{N}$. This implies that it is possible to synthesize terms of the form $\epsilon^{2k+1}\Omega_x$ through successive Lie bracketing. With available generators of rotation $\{\epsilon\Omega_x, \epsilon^3\Omega_x, \dots, \epsilon^{2n+1}\Omega_x\}$ can produce rotations,

$$R_x(\epsilon) = \exp(c_0\epsilon\Omega_x) \exp(c_1\epsilon^3\Omega_x) \dots \exp(c_n\epsilon^{2n+1}\Omega_x) = \exp\left(\sum_{k=0}^n c_k \epsilon^{2k+1} \Omega_x\right).$$

We can then choose the order of approximation n and the coefficients $\{c_k\}$ so that $\sum_{k=0}^n c_k \epsilon^{2k+1} \approx \Theta_x(\epsilon)$. By approximating any odd function in this manner we can generate a rotation,

$$R_x(\epsilon) \approx \exp\{\Theta_x(\epsilon)\Omega_x\}.$$

about the x -axis. Similarly, we can generate an approximate rotation around the y -axis with angle $\Theta_y(\epsilon)$. Therefore, using Euler angle decomposition with angles (α, β, γ) we can synthesize any arbitrary rotation in $\text{SO}(3)$,

$$R(\epsilon) = \exp\{\alpha(\epsilon)\Omega_x\} \exp\{\beta(\epsilon)\Omega_y\} \exp\{\gamma(\epsilon)\Omega_x\},$$

provided $R(\epsilon)$ is a continuous function in ϵ , by approximating the functions $\alpha(\epsilon)$, $\beta(\epsilon)$, and $\gamma(\epsilon)$ as done above. Since we can achieve any desired rotation with arbitrary accuracy, through the choice of n , this system is controllable.

This is in fact a constructive proof for controllability of the Bloch equations; however, generating back-and-forth evolutions is inefficient in practice. Creating methods to design efficient compensating controls is then of particular importance. The involvement of polynomial approximation in the proof of controllability suggests that polynomials should be incorporated in the methods developed to control and optimize these systems. In the real world application of the Bloch equations, there are bounds on, for example, the available control amplitudes. In NMR the spectrometers have sensitive probes and in MRI there are limits on the allowable applied fields that can be used with human patients. In dissipative quantum systems the system state is not constant and so the desired target state is unknown but should be maximized. These considerations lead us to consider the important practical question of optimal control of ensembles.

2.1.2 Controllability of Time-Varying Linear Ensemble Systems

In particular, time-varying linear systems of the following form have been studied,

$$\frac{d}{dt}x(t, s) = A(t, s)x(t, s) + B(t, s)u(t),$$

where the state $x \in \mathbb{R}^n$ is indexed by a parameter $s \in \mathcal{S}$ that shows variation, with A and B time-varying matrices of appropriate size that are also indexed by s . The single open-loop control $u \in \mathbb{R}^m$ is used to steer the ensemble of systems from an initial state (or distribution) to another state (or distribution). The state of such a system, starting at $x(0, s)$, at any time t can be expressed by the variation of constants formula,

$$x(t, s) = \Phi(t, 0, s)x(0, s) + \int_0^t \Phi(t, \tau, s)B(\tau, s)u(\tau) d\tau,$$

where $\Phi(t, 0, s)$ is the transition matrix associated with the linear system and can be computed by the Peano-Baker series if A is bounded [6]. If we let $\xi(s) =$

$\Phi(0, t, s)x_d(s) - x(0, s)$, where $x_d(s)$ is the desired final state, then the variation of constants formula above becomes,

$$\xi(s) = \int_0^t \Phi(0, \tau, s)B(\tau, s)u(\tau) d\tau. \quad (2.5)$$

It has been shown that the necessary and sufficient conditions for ensemble controllability are related to the Fredholm integral equation of the first kind of the input-to-state operator [37],

$$(Lu)(s) = \int_0^T \Phi(0, t, s)B(t, s)u(t) dt, \quad (2.6)$$

which is found by letting $t = T$ in (2.5). For controllability, we require $(Lu)(s) = \xi(s)$, where $\xi(s)$ is known as long as the final state $x_d(T)$ can be specified. The conditions on the singular system of this operator guarantee controllability of this class of ensemble systems.

2.2 Optimal Control

Optimal control merges the two individually challenging areas of control and optimization to rigorously characterize and solve optimization problems based on dynamical systems. Early versions of optimal control can be seen as far back as Bernoulli's Brachystochrone problem in 1696 with the birth of the calculus of variations. Subsequent associated ideas, including the Euler-Lagrange equations of motion, pushed further towards the concept of optimal control [64]. A flourish of work starting in the mid-twentieth century marked the true era of optimal control including Pontryagin's maximum principle [52].

Common amongst all optimal control problems is a dynamical system of the general form,

$$\frac{d}{dt}x(t) = f(t, x(t), u(t)),$$

and a specified initial state $x(0) = x_0$. Beyond these elements, the nature and composition of an optimal control problem can vary depending upon the application. In some cases a terminal state constraint can be given, i.e., $x(T) = x_d$, where x_d is the

desired final state. Whether or not such a terminal constraint is given, a function of merit is defined to evaluate candidate control and state trajectory pairs. This cost functional incorporates a terminal cost to be evaluated at the final time, T , and a running cost that evaluates the time history of the states and controls. Additional path constraints can also be incorporated to place variable bounds on the solution. Assembling these individual pieces yields an optimal control problem of standard Bolza form,

$$\begin{aligned}
\min \quad & \varphi(T, x(T)) + \int_0^T \mathcal{L}(x(t), u(t)) \, dt \\
\text{s.t.} \quad & \frac{d}{dt}x(t) = f(t, x(t), u(t)), \\
& e(x(0), x(T)) = 0, \\
& g(x(t), u(t)) \leq 0,
\end{aligned} \tag{2.7}$$

where we can minimize (or maximize) the terminal and running costs, φ and \mathcal{L} , respectively, subject to the dynamics; endpoint constraints given by e ; and additional path constraints given by g [9].

2.2.1 Solving Optimal Control Problems

As described above, optimal control is a general framework to pose optimizations of dynamical systems. A variety of methods have been developed and employed to solve such problems. Here we give a brief overview of this area to provide context for the numerical method presented in Chapter 3.

Maximum Principle

The maximum principle states the necessary conditions for a candidate control, $u(t)$, to be optimal with respect to the optimal control problem in (2.7). It is stated as follows.

Theorem 1 (Maximum Principle [52]). *Suppose the controlled trajectory (x_*, u_*) defined over the interval $[0, T]$ is optimal and the control u_* is piecewise continuous.*

Then there exist a constant $\lambda_0 \geq 0$ and a covector $\lambda : [0, T] \rightarrow (\mathbb{R}^n)^*$, the so-called adjoint variable, such that the following conditions are satisfied:

1. **Nontriviality** of the conditions: $(\lambda_0, \lambda(t)) \neq 0, \forall t \in [0, T]$;
2. **Adjoint equation:** the adjoint variable λ is a solution to the time-varying linear differential equation

$$\dot{\lambda}(t) = -\lambda_0 L_x(t, x_*(t), u_*(t)) - \lambda(t) f_x(t, x_*(t), u_*(t)); \quad (2.8)$$

3. **Minimum condition:** everywhere in $[0, T]$ we have

$$H(t, \lambda_0, \lambda(t), x_*(t), u_*(t)) = \min_{v \in U} H(t, \lambda_0, \lambda(t), x_*(t), v). \quad (2.9)$$

If the Lagrangian \mathcal{L} and the dynamics f are continuously differentiable in t , then the function $h : t \rightarrow H(t, \lambda_0, \lambda(t), x_*(t), u_*(t))$ is continuously differentiable with derivative given by

$$\dot{h}(t) = H_t(t, \lambda_0, \lambda(t), x_*(t), u_*(t)). \quad (2.10)$$

4. **Transversality condition:** at the endpoint of the controlled trajectory the covector

$$(H + \lambda_0 \varphi_t, -\lambda + \lambda_0 \varphi_x)$$

is orthogonal to the terminal constraint N , i.e., there exists a covector $\nu \in (\mathbb{R}^{n+1-k})^*$ such that

$$H + \lambda_0 \varphi_t + \nu D_t \Psi = 0, \quad \lambda = \lambda_0 \varphi_x + \nu D_x \Psi \quad \text{at } (T, x_*(T)). \quad (2.11)$$

Here we label the state space and control set M and U , respectively. We define a Hamiltonian,

$$H(t, \lambda_0, \lambda, x, u) = \lambda f(t, x, u) + \lambda_0 \mathcal{L}(t, x, u),$$

terminal constraints as a k -dimensional embedded submanifold,

$$N = \{(t, x) \in \mathbb{R} \times M : \Psi(t, x) = 0\},$$

and D in (2.11) is the respective matrix of partial derivatives.

In principle, the goal of the maximum principle is to eliminate the optimization by incorporating the cost into a two point boundary value problem on both the state equation dynamics and the adjoint dynamics in (2.8). By solving these equations simultaneously we can identify candidate extremals that satisfy both the dynamics and the adjoint equation, thereby satisfying the first order necessary conditions for optimality.

Computational Methods

While optimal control gives a mathematically rigorous way to frame optimization problems on dynamical systems and existing results, like the maximum principle, can give the necessary criteria for an optimal solution, using this machinery is largely nontrivial. Analytic work in this area can be challenging for systems of even one or two dimensions. Due to these difficulties, significant interest has been garnered in the area of computational optimal control methods. Numerical methods for optimal control can be classified either as a direct or indirect method. Although these methods can use the same numerical methods to solve them, they differ in implementation and convergence.

An indirect method focuses on finding an approximate solution to the two point boundary value problem from the necessary conditions, as given by the maximum principle. The advantages to solving this set of coupled state and adjoint equations instead of the optimal control problem include increased accuracy and the guarantee that the solution satisfies the necessary optimality conditions. However, these conditions must be derived analytically, which poses a great challenge to nonlinear and high dimension problems. There are also a number of cases in which the maximum principle does not give sufficient information to specify the endpoint constraints on the two point boundary value problem, such as when the target state is a single point and the Transversality condition does not yield additional information to set the endpoint constraints of the adjoint equation. In these cases, values for the endpoints must be guessed and evaluated, which requires another level of an iterative scheme.

In comparison, direct methods approach the original optimization problem in (2.7). While they lack some of the accuracy inherent in indirect methods, they require little additional information regarding initial guesses and do not need iterative schemes to test various endpoint values. The state variables are physically relevant to the problem, unlike the adjoint variable in the necessary conditions. As mentioned above, similar methods - such as shooting, finite difference, collocation, etc. - can be used with direct or indirect methods. The easy interpretation and implementation of direct methods make them prime candidates for study. Moreover, moving toward parameterized systems in which deriving the necessary conditions is generally infeasible, we consider a direct method in this work to build a computational scheme for solving optimal ensemble control problems.

2.3 Optimal Ensemble Control

The time-evolution of ensemble systems requires a slightly more general form for the dynamics than given in (2.7). In particular, the ensemble dynamics (2.1) includes an additional dimension of continuity in the parameter over the specified interval of variation. The following formulation for optimal ensemble control must accommodate this extra dimension,

$$\begin{aligned}
\min \quad & \int_{\mathcal{S}} \left[\varphi(T, x(T, s)) + \int_0^T \mathcal{L}(x(t, s), u(t)) dt \right] ds, \\
\text{s.t.} \quad & \frac{d}{dt} x(t, s) = F(t, s, x(t, s), u(t)), \\
& e(x(0, s), x(T, s)) = 0, \\
& g(x(t, s), u(t)) \leq 0,
\end{aligned} \tag{2.12}$$

where the cost functional includes another integration in the parameter, s . The cost is attempted to be highly general. We implicitly assume that the terminal and running costs can include weighting functions of s , i.e., $\varphi = \varphi(s, T, x(T, s))$ and $\mathcal{L} = \mathcal{L}(s, t, x(t, s), u(t))$. In this way the cost can incorporate any desired profile in the parameter domain \mathcal{S} .

An optimal nonlinear control problem of this form is, in general, analytically intractable. Computational methods are then required to solve such exceedingly complex optimal ensemble control problems. The idea from previous work in ensemble control of quantum systems that constructing appropriate polynomials is a key tool in characterizing the controllability of ensemble systems of interest motivates the use of polynomials within the computational framework [39].

Chapter 3

Pseudospectral Method for Optimal Control

As we have seen in the preceding chapters, the challenge in solving general nonlinear optimal control problems and, even more so, optimal ensemble control problems motivates the investigation of robust and flexible numerical methods. In this work I adapt a Legendre pseudospectral method, which is a spectral collocation method using Legendre polynomials as basis functions for the expansion of the unknown variables. Spectral methods were originally developed to solve problems in fluid dynamics and since then pseudospectral methods have been successfully applied to optimal control problems in many areas of science and engineering [54, 16, 56] since the introduction of the concept in the mid 1990s [14]. The fact that ensemble dynamics share some similar characteristics with partial differential equations supports that a pseudospectral method is well suited as a numerical method for optimal ensemble control.

3.1 Spectral Methods

Spectral methods are numerical techniques that rely on the expansion of a function in terms of an infinite sequence of orthogonal functions, i.e. $x = \sum_{k=0}^{\infty} \hat{x}_k \phi_k$ where $\langle \phi_k, \phi_j \rangle = \delta_{kj}$. These methods are characterized by spectral accuracy, which means that the k^{th} coefficient, \hat{x}_k , decays faster than k^{-n} , $\forall n \geq 1$, for smooth functions [10]. The well-known Fourier series expansion for periodic functions is possibly the most notable example. Fourier series are limited to expression of periodic functions due to the choice of basis (sometimes called trial) functions, ϕ_k , the trigonometric functions.

Through a different choice of orthogonal functions, namely those of Jacobi type, we can extend spectral methods to non-periodic functions. For our purposes, spectral methods are a natural numerical framework for dealing with optimal control problems, because they are able to transform an ordinary or partial differential equation into an algebraic equation through relatively simple recursion relations for the derivatives of the basis functions.

There are three different approaches to spectral methods. The basis functions express the approximation to the solution of the differential equation; however, another set of functions, weight (sometimes called test) functions - the \hat{x}_k , are needed to ensure that the differential equation and boundary conditions are satisfied. How the weight functions are selected separates the various spectral methods and is directly related to the residual - the error between the truncated expansion and the exact function [4]. For optimal control, we are interested in differential equations that govern dynamics, in general, $\dot{x} = f(x)$. The residual for a truncated expansion $x^N = \sum_{k=0}^N \hat{x}_k \phi_k$ is

$$R(x) = \frac{d}{dt}x^N - f(x^N).$$

Collocation methods require the residual to be zero at specific physically-relevant points whereas tau and Galerkin methods enforce the differential equation by requiring that the residual is orthogonal to as many of the basis functions as possible. Collocation methods offer a distinct advantage for differential equations involving variable coefficients or nonlinear terms, as the computation involves products of numbers - the values at the collocation nodes - rather than products of expansions as in Tau or Galerkin methods [18]. It is from this difference that the collocation spectral method gets the “pseudo”-spectral name.

3.1.1 Comparison to Finite Element Methods

Whereas collocation, Galerkin, and tau spectral methods differ in their choice of test functions, spectral methods differ from finite element methods by their choice of basis functions. The idea of approximating a function on an interval with a polynomial is

common to both, however, finite elements select basis functions to be local polynomials on small sub-intervals, generally of much lower order [10]. In order to increase accuracy the three major types of finite element methods either (“h-refinement”) subdivide each sub-interval further to improve resolution, (“r-refinement”) only subdivide sub-intervals where the gradient is steep, or (“p-refinement”) increase the degree of the polynomial on each interval [4]. The major advantage of finite element methods is the ability to chop up complex multidimensional structures, such as engineered surfaces, into sub-intervals, which generally comes at the cost of accuracy. Therefore, the use of spectral versus finite element tools is largely determined by the application. In the case of modeling the compression dynamics of an automobile during a collision, the geometry of the structure and the small scale of fractures and distortions would be more appropriate for finite element methods. In the case of optimal control, in which solutions are relatively smooth (or piecewise smooth), global polynomial expansions of reasonable order can be expected to not only work well but out perform finite element methods.

3.2 Pseudospectral Method

The overarching goal of the pseudospectral method is to convert the continuous optimal control problem in (2.7) into a constrained algebraic optimization problem, which can be solved by efficient numerical nonlinear optimization solvers. To provide a context for understanding this work, I first present a review of the pseudospectral method for optimal control, i.e., without variation, and then provide an extension to consider the optimal ensemble control problem.

Pseudospectral discretization methods use expansions of orthogonal polynomials (see Appendix A.1) to approximate the states of the system and thereby inherit the spectral accuracy characteristic of orthogonal polynomial expansions [10]. Through special recursive properties (see Appendix A.2), derivatives of these orthogonal polynomials can be expressed in terms of the polynomials themselves, making it possible to accurately approximate the differential equation that describes the dynamics with an algebraic relation imposed at a small number of discretization points. An appropriate choice of these discretization points, or nodes, facilitates the approximation of the

states as well as ensures accurate numerical integration through Gaussian quadrature. As we elaborate on these ideas below, a majority of the derivations and justifications behind the formulas in this section have been placed in Appendix A to make this important section more readable.

We first shift the original problem from the time domain $t \in [0, T]$ to the rescaled domain $\tilde{t} \in [-1, 1]$ on which the orthogonal polynomials are defined with a simple affine transformation

$$\tilde{t} = \frac{2t - T}{T}.$$

In this dissertation, the variable t will be used interchangeably either on the $[0, T]$ or $[-1, 1]$ interval and the choice will be set by context. Our choice of the Legendre orthogonal polynomial family suggests we compute the integral term of the cost function using Legendre-Gauss-Lobatto (LGL) quadrature, in which the integral is approximated by a summation of the integrand evaluated at specific set of nodes,

$$\int_{-1}^1 f(t) dt \approx \sum_{i=1}^N f(t_i) w_i, \quad w_i = \int_{-1}^1 \ell_i(t) dt, \quad (3.1)$$

where N is the order of polynomial approximation, w_i are discrete weights, and $\ell_i(t)$ is the i^{th} Lagrange polynomial, discussed below and in Appendix A.3 [4]. Lobatto in LGL refers to the inclusion of the endpoints as nodes, which is necessary in discretizing optimal control problems in order to enforce initial and terminal conditions. In particular, if the integrand $f \in \mathbb{P}_{2N-1}$ and the nodes $t_i \in \Gamma^{\text{LGL}}$, the integral approximation is exact, where \mathbb{P}_{2N-1} denotes the set of polynomials of degree $2N - 1$ or less and where $\Gamma^{\text{LGL}} = \{t_i : L'_N(t)|_{t_i} = 0, i = 1, \dots, N - 1\} \cup \{-1, 1\}$ are the $N + 1$ LGL nodes determined by the derivative of the N^{th} order Legendre polynomial, $L'_N(t)$, and the interval endpoints [10].

LGL quadrature requires we know the integrand values at the LGL nodes - $f(t_i)$ in (3.1); however, the N^{th} order Legendre expansions

$$x(t) \approx P_N x(t) = \sum_{k=0}^N \tilde{x}_k L_k(t), \quad (3.2)$$

$$u(t) \approx P_N u(t) = \sum_{k=0}^N \tilde{u}_k L_k(t), \quad (3.3)$$

do not directly give us a way to discretize the states and controls at these nodes, i.e., the expansions coefficients \tilde{x}_k and \tilde{u}_k have no direct physical meaning. To overcome this, we approximate these Legendre expansions with interpolating polynomials, which, by definition, are equal to the Legendre expansions at the interpolation nodes. Because any interpolating polynomial can be represented by Lagrange polynomials we can represent the state and control as,

$$P_N x(t) \approx I_N x(t) = \sum_{k=0}^N \bar{x}_k \ell_k(t), \quad (3.4)$$

$$P_N u(t) \approx I_N u(t) = \sum_{k=0}^N \bar{u}_k \ell_k(t), \quad (3.5)$$

where the coefficients \bar{x}_k and \bar{u}_k are the values of the state and control Legendre expansions evaluated at the k^{th} interpolation node, respectively, i.e., $P_N x(t_k) = I_N x(t_k) = \bar{x}_k$ and $P_N u(t_k) = I_N u(t_k) = \bar{u}_k$. The coefficients have this property because the k^{th} Lagrange polynomial is characterized by taking unit value at the k^{th} interpolation node and zero value at all other nodes such that $\ell_k(t_i) = \delta_{ki}$, where δ_{ki} is the Kronecker delta function [65]. Using this second approximation we can compute the integrand of the cost function integral at the LGL nodes and \bar{x}_k and \bar{u}_k become the decision variables of the subsequent nonlinear programming problem.

Furthermore, the selection of LGL nodes, which are non-uniform on $[-1, 1]$ with quadratic spacing towards the endpoints, as interpolation nodes suppresses the spurious oscillations between nodes that is present when using uniformly spaced nodes, called the Runge phenomena [18]. It can be shown that the LGL interpolation nodes are close to optimal (see Appendix A.4) [62]. The LGL nodes permit us to rewrite the Lagrange polynomials in terms of the Legendre polynomials, which is critical to inherit the special derivative and spectral accuracy properties of the orthogonal polynomials despite using Lagrange interpolating polynomials. Given $t_k \in \Gamma^{\text{LGL}}$, we can express the Lagrange polynomial as (see Appendix A.5),

$$\ell_k(t) = \frac{1}{N(N+1)L_N(t_k)} \frac{(t^2 - 1)L'_N(t)}{t - t_k}. \quad (3.6)$$

The derivative of (3.4) at $t_i \in \Gamma^{\text{LGL}}$ is then,

$$\begin{aligned} \frac{d}{dt} I_N x(t_i) &= \sum_{k=0}^N \bar{x}_k \dot{\ell}_k(t_i) = \sum_{k=0}^N D_{ik} \bar{x}_k \\ &= D_{i0} \bar{x}_0 + D_{i1} \bar{x}_1 + \cdots + D_{iN} \bar{x}_N, \end{aligned} \quad (3.7)$$

where D is the constant matrix with elements (see Appendix A.6),

$$D_{ik} = \begin{cases} \frac{L_N(t_i)}{L_N(t_k)} \frac{1}{t_i - t_k} & i \neq k \\ -\frac{N(N+1)}{4} & i = k = 0 \\ \frac{N(N+1)}{4} & i = k = N \\ 0 & \text{otherwise.} \end{cases} \quad (3.8)$$

We have now effectively discretized all parts of the original optimal control problem. The problem in (2.7) can now be written as,

$$\begin{aligned} \min \quad & \varphi(T, \bar{x}_N) + \frac{T}{2} \sum_{i=0}^N \mathcal{L}(\bar{x}_i, \bar{u}_i) w_i \\ \text{s.t.} \quad & \sum_{k=0}^N D_{ik} \bar{x}_k = \frac{T}{2} f(\bar{x}_i, \bar{u}_i), \\ & e(\bar{x}_0, \bar{x}_N) = 0, \\ & g(\bar{x}_i, \bar{u}_i) \leq 0, \quad \forall i \in \{0, 1, \dots, N\}. \end{aligned} \quad (3.9)$$

Notice that the second and third lines are equality constraints reflecting the dynamics and endpoint conditions, respectively, and the last line is an inequality constraint reflecting the path constraints.

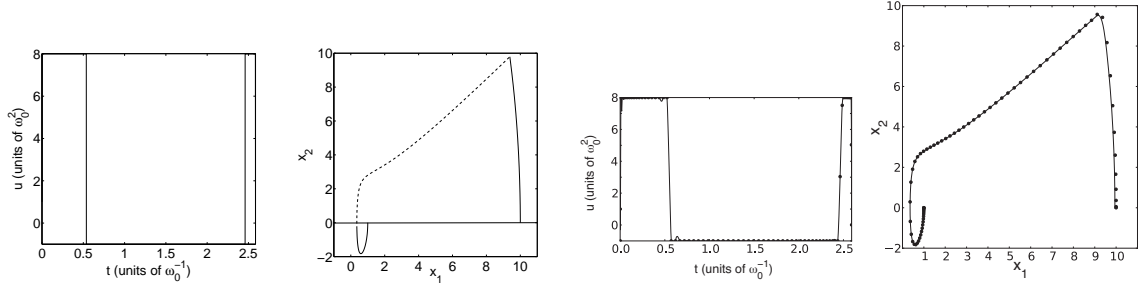


Figure 3.1: The time-optimal frictionless atom cooling problem in (3.10) is efficiently solved by the pseudospectral method. The dots in the two right-most figures indicates the 71 interpolation nodes used to approximate the original problem.

3.3 Example from Quantum Optics

Now that we have laid the foundation for the pseudospectral method for optimal control, it is worth evaluating its performance on a pertinent example. Frictionless atom cooling in a harmonic trap refers to lowering the harmonic frequency of the trap to a desired value, while keeping the populations of the initial and final levels the same. In order ensure that the levels are not changed, friction and heat must not be generated throughout this process. There are adiabatic methods to achieve this transfer, however, these are typically much slower than necessary. We consider the specific problem here to design the time optimal control, where we allow the trap to become an expulsive potential, corresponding to imaginary frequencies, for a short period of time. The dynamics of this problem can be transformed into the following normalized form [63],

$$\begin{aligned}\dot{x}_1 &= x_2, \\ \dot{x}_2 &= -ux_1 + \frac{1}{x_1^3}.\end{aligned}\tag{3.10}$$

We seek to find the control $-1 \leq u(t) \leq 8$, with $u(0) = 1$ and $u(T) = 1/\gamma^4$, that steers (3.10) from $(x_1(0), x_2(0)) = (1, 0)$ to $(x_1(T), x_2(T)) = (\gamma, 0)$ minimizing the final time, T , where $\gamma = \sqrt{\omega_0/\omega_f} = 10$ quantifies the desired cooling from a natural frequency ω_0 to a desired frequency ω_f .

This problem has been solved analytically and the derived control and corresponding trajectories are shown in the first two frames of Figure 3.1 [63]. The pseudospectral method effectively transforms the continuous time-optimal control problem into a nonlinear programming problem and finds a close approximation of the optimal control and state trajectory. The optimized control and trajectory are also shown in Figure 3.1 with the interpolation nodes overlaid as dots. The cubic term in the dynamics and the high resolution needed for “bang-bang” type control required a level of discretization around $N = 71$.

3.4 Multidimensional Extension

The ensemble optimal control problem in (2.12) includes another dimension of continuity in the parameter domain, $s \in \mathcal{S} \subset \mathbb{R}^d$, which must be discretized (or sampled) to fit within the constrained optimization method. To reduce the complexity of notation we first consider a single parameter variation, i.e., $d = 1$ and $\mathcal{S} = [\underline{s}, \bar{s}]$, however, it is straightforward to extend this to higher dimensions.

Consider the ensemble extension of the interpolation approximation in (3.4),

$$\begin{aligned} x(t, s) &\approx I_{N \times N_s} x(t, s) = \sum_{k=0}^N \bar{x}_k(s) \ell_k(t) \\ &\approx \sum_{k=0}^N \left(\sum_{r=0}^{N_s} \bar{x}_{kr} \ell_r(s) \right) \ell_k(t), \end{aligned} \quad (3.11)$$

and the ensemble extension of the approximate derivative from (3.7) at $t_i \in \Gamma^{\text{LGL}}$ and $s_j \in \Gamma_{N_s}^{\text{LGL}}$,

$$\begin{aligned} \frac{d}{dt} I_{N \times N_s} x(t_i, s_j) &= \sum_{k=0}^N D_{ik} \left(\sum_{r=0}^{N_s} \bar{x}_{kr} \ell_r(s_j) \right) \\ &= \sum_{k=0}^N D_{ik} \bar{x}_{kj}, \end{aligned} \quad (3.12)$$

where $\bar{x}_{kj} = x(t_k, s_j)$. In (3.11) and (3.12) we have effectively used a two dimensional interpolating grid at the $N + 1$ and $N_s + 1$ LGL nodes in time and the parameter,

respectively. Using (3.11), (3.12), in conjunction with the LGL quadrature rule, we summarize the ensemble pseudospectral discretization of the ensemble optimal control problem as

$$\begin{aligned}
\min \quad & \frac{\bar{s} - s}{2} \sum_{r=0}^{N_s} \left[\varphi(T, \bar{x}_{Nr}) + \frac{T}{2} \sum_{i=0}^N \mathcal{L}(\bar{x}_{ir}, \bar{u}_{ir}) w_i^N \right] w_r^{N_s} \\
\text{s.t.} \quad & \sum_{k=0}^N D_{ik} \bar{x}_{kr} = \frac{T}{2} F(\bar{x}_{ir}, \bar{u}_i), \\
& e(\bar{x}_{0r}, \bar{x}_{Nr}) = 0, \\
& g(\bar{x}_{ir}, \bar{u}_{ir}) \leq 0, \quad \forall \begin{matrix} i \in \{0, 1, \dots, N\} \\ r \in \{0, 1, \dots, N_s\} \end{matrix},
\end{aligned} \tag{3.13}$$

where w^N and w^{N_s} are the LGL quadrature weights corresponding to polynomial approximations of order N and N_s respectively.

It is straightforward to extend this interpolating structure to accommodate parameter spaces of higher dimension, $\mathbf{s} = (s_1, s_2, \dots, s_d)' \in \mathcal{S} \subset \mathbb{R}^d, d > 1$. In this general case,

$$\begin{aligned}
x(t, \mathbf{s}) &\approx I_{N \times N_{s_1} \times \dots \times N_{s_d}} x(t, \mathbf{s}) = \sum_{k=0}^N \bar{x}_k(\mathbf{s}) \ell_k(t) \\
&= \sum_{k=0}^N \sum_{r_1=0}^{N_{s_1}} \dots \sum_{r_d=0}^{N_{s_d}} \bar{x}_{kr_1 \dots r_d} \ell_{r_d}(s_d) \dots \ell_{r_1}(s_1) \ell_k(t).
\end{aligned}$$

3.5 Implementation

As a significant portion of this dissertation deals with numerical computation, it is appropriate to highlight certain aspects of the implementation of the method and discuss properties of a practical nature. As mentioned in preceding sections of this chapter, the objective of the pseudospectral and multidimensional pseudospectral method is to approximate a continuous-time optimal control problem by a nonlinear programming problem. The advantage in doing this is that by discretizing the continuous dynamics and converting the state space from a function space to a vector space, we can leverage decades of work in algorithm development for nonlinear optimization, such as all types of gradient methods, simulated annealing, genetic algorithms, and

many others. Most modern nonlinear solvers include multiple of these tools, which when used in concert allow for some level of confidence regarding avoidance of local minima and other common shortcomings of individual numerical solvers.

3.5.1 Visualization

To make the translation of the optimal control problem into the discretized optimization more concrete, this section expands on several of the key steps within the pseudospectral approximations to make them more immediately accessible. The original problem has unknown functions $x(t) \in \mathbb{R}^n$ and $u(t) \in \mathbb{R}^m$ and a possibly unknown terminal time T . Approximating the states and controls with polynomials of order N on the nodes Γ_N^{LGL} as in (3.4) creates interpolating polynomials with coefficients \bar{x}_i and \bar{u}_i

$$\begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \\ \hline u_1(t) \\ \vdots \\ u_m(t) \end{bmatrix} \Rightarrow \begin{bmatrix} \bar{x}_{10} & \bar{x}_{11} & \cdots & \bar{x}_{1N} \\ \vdots & & & \vdots \\ \bar{x}_{n0} & \bar{x}_{n1} & \cdots & \bar{x}_{nN} \\ \hline \bar{u}_{10} & \bar{u}_{11} & \cdots & \bar{u}_{1N} \\ \vdots & & & \vdots \\ \bar{u}_{m0} & \bar{u}_{m1} & \cdots & \bar{u}_{mN} \end{bmatrix}$$

which become the decision variables of the discretized optimization in addition to the terminal time T . In the ensemble case, when $x(t, s) \in \mathbb{R}^n$, the decision variables compose a $(d+2)$ -dimensional matrix, where $s \in \mathbb{R}^d$. In the discretized optimization, similar to the original optimal control problem, the solver then selects the values \bar{u}_{ij} subject to the bounds placed on the control and any specific path constraints. Setting the discretized optimization apart from the original, the state coefficients \bar{x}_{ij} are also treated as decision variables and the solver chooses values for these coefficients subject to bounds on the state, specific path constraints, and also the equality constraints that impose the dynamics as in (3.7).

3.5.2 Limitations

While the pseudospectral method is a highly robust and flexible method, there are some limitations of the method that should be mentioned. In particular there are discontinuous or highly varying solutions that it will not find, due to the coarse discretization and polynomial approximation. Although it will not find this class of controls and the corresponding states, in practical applications it is of great benefit to develop solutions that closely approximate the discontinuities while remaining smooth. As with all optimization methods, it is susceptible to local minima. We can, especially if the optimal (or anticipated) cost is known, impose a “performance” constraint that requires the solver to continue searching until it achieves this specified level.

Cost function selection is highly influential on the optimization outcome and similar cost function choices can, at times, lead to different results. It is important to fully understand the dynamics of the system so as to avoid, for example, singularities in the state search space. Finally, adjusting and testing various parameter values (e.g. N , N_s) is key to finding the right tradeoff between speed and accuracy.

3.5.3 Programming Environments & Languages

The “pseudospectral method” is a collection of mathematical approximations, which means the method is independent of the exact implementation of it. The fundamental requirements are to have some ability for the necessary linear algebra as well as the definitions of the Lagrange and Legendre polynomials. With these tools, scripts can be written to calculate the LGL nodes and the coefficients in the differentiation matrix, D . Therefore, any programming language with access to these types of libraries is suitable for use with the pseudospectral method. In fact, there are a variety of commercial and open source packages available for various forms of the pseudospectral method for optimal control. These packaged solutions are mainly written in MATLAB, however, C, Python, Java, and any other major programming language are fully viable. The ensemble extension, although straightforward, requires a level of flexibility that is generally not found in these commercial solutions, which is why in this research I built custom software to perform the pseudospectral discretization.

Although I began this work prototyping in MATLAB, there is tremendous benefit to working with a highly specialized optimization language like AMPL. AMPL is both a program and a syntax structure that facilitates the succinct formulation of a vector space optimization problem. The patterns of the pseudospectral approximations, and further the patterns inherent in ensemble control problems, make it simple to write the nonlinear programming problem in this syntax. Beyond this ease of implementation, AMPL offers a useful presolve and plug-and-play features. A presolve is done by AMPL to eliminate any redundant or check for conflicting constraints. This eliminates unnecessary degrees of freedom, which simplifies the problem before it is handed to a numerical solver. AMPL is a widely accepted optimization interface and all mainstream numerical solvers create versions that work immediately with AMPL. This provides a wealth of options when selecting a single solver for the optimization as compared to MATLAB or other languages, where solver support is less immediate. It should be noted that another language is required to compute the D matrix given a specific N . In this work, I wrote several Python scripts to automate the process of writing the AMPL code and processing the AMPL output, e.g. plotting.

Because of the stand-alone nature and text file input-output of AMPL, it is well suited to be used as the backend on an optimization webservice. There is a federally funded optimization initiative called NEOS that uses a similar structure. During my work, both to facilitate my own research and provide collaborators access to the tools I developed, I created a Django webservice (a web framework with a Python backend) that could be used to submit automated optimization “jobs” and keep track of them. This is the foundation that will allow the research group to eventually share the power and flexibility of the pseudospectral method with a much broader audience. In particular, scientists lacking the knowledge or resources to implement our method can likewise submit jobs for their specific application and get results by email. Giving such scientists easy access to tools that can significantly impact their physical, chemical, and biological experiments, this webservice has the potential to revolutionize these fields.

Two examples are included in Appendix B to illustrate the short and efficient code used with AMPL. The full formulation of these examples is given in the appendix, however, a better understanding of them will come after reading the next chapter.

Chapter 4

Optimal Pulse Design for NMR & MRI

The consideration of variations in practical NMR and MRI led to the development of ensemble control theory, which established the potential to construct pulses that would compensate for the dispersion in the inhomogeneous dynamics. Realistic concerns over energy use and goals to maximize the performance, efficiency, and resolution in quantum applications then guides us to cast pulse design problems as optimal ensemble control problems. The complexity inherent in these pulse designs requires us to develop computational techniques, such as the multidimensional pseudospectral method discussed in Chapter 3. This chapter presents solutions to pulse design problems generated by the multidimensional pseudospectral method and highlights specific features of the method that make it effective on this class of problems.

4.1 Pulse Design

Compelling applications for quantum control have received particular attention and have motivated seminal studies in wide-ranging areas from coherent spectroscopy and MRI to quantum optics. Designing and implementing time-varying excitations (rf pulses) to manipulate complex dynamics of a large quantum ensemble on the order of Avogadro's number is a longstanding problem and an indispensable step that enables every application of quantum control [44]. For example, magnetic resonance applications often suffer from imperfections such as inhomogeneity in the static magnetic field (B_0 inhomogeneity) and in the applied rf field (rf inhomogeneity). In addition,

there is dispersion in the Larmor frequency of spins due to chemical shifts. A good pulse design strategy must be robust to these effects and such variations need to be considered in the modeling and pulse design stages in order to match theoretical predictions to experimental outcomes. As difficult experiments with more demanding performance specifications have emerged, the complexity of finding optimal pulse sequences has drastically increased. For example, as high-field NMR spectrometers are increasingly more accessible and required, broadband excitation pulses are needed to cover a wide ^{13}C chemical-shift range, e.g., up to 40 kHz. In addition, to design excitation and inversion pulses that are practical for a typical NMR spectrometer, methods must accommodate realistic maximum rf power and pulse duration while accomplishing the desired spin evolution. Such limitations and imperfections cause a substantial increase in the complexity of the pulse design problem.

From early work using physical intuition [25, 11] to modern methods like composite pulses [35, 60], an enormous body of pulse sequence design techniques has been proposed over 50 years [2, 12] and the process of innovation is ongoing. Highly customized methods, however, have limited scope, such as the Shinnar-Le Roux algorithm which is robust to Larmor dispersion, but not able to compensate for rf inhomogeneity [50]. For relatively simple cases, theoretical methods, such as average Hamiltonian theory, provide intuitive guidelines for constructing pulse sequences [23]. Heuristic numerical optimization methods have been used extensively for the design of single and multiple pulses in a pulse sequence [59]. However, these approaches have a number of shortcomings such as slow convergence rates and being easily trapped into local optima. In recent years, there have been attempts to look at pulse design problems from a control theory perspective [13, 53, 61, 28]. In particular, state-of-the-art methods such as gradient ascent and Krotov algorithms are based on principles of optimal control theory [31, 46] and have been used successfully to design broadband and relaxation-optimized pulses which maximize the performance of quantum systems in the presence of relaxation [33, 29, 19]. These algorithms, while effective, rely on intensive computations, as for system propagators and gradients, as well as a large number of discretizations in the time domain over which to evolve the system.

To overcome these defects we provide a systematic framework for optimal pulse design in quantum control. We consider a general mathematical model for pulse design as an optimal control problem of a continuum of bilinear systems. Employing Lie

algebra tools from prior theoretical work [39], we have shown the control problem of pulse design can be mapped to a problem of polynomial approximation. This new notion guides us to develop a unified computational method for optimal pulse design based on multidimensional pseudospectral approximations, by which a continuous-time optimal control problem of pulse design can be discretized to a constrained optimization problem using interpolating polynomials [40, 41, 57, 58, 56, 42].

4.2 Quantum Dynamics

The dynamics of a quantum system are given by the time-evolution of its density matrix. We consider here general dynamics in which the system may have interaction with the environment that leads to dissipation in the system state. Under the Markovian approximation, where the environment is modeled as an infinite thermostat which has constant state, the evolution of the density matrix can be written in Lindblad form in terms of the system Hamiltonian $H(t)$ and superoperator $L(\cdot)$ which model the unitary and nonunitary dynamics [43], respectively,

$$\frac{d}{dt}\rho = -i[H(t), \rho] - L(\rho), \quad (\hbar = 1).$$

The expression of the Hamiltonian has components corresponding to the free evolution Hamiltonian, H_f , and the control Hamiltonians H_i ,

$$H(t) = H_f + \sum_{i=1}^m u_i(t)H_i,$$

where $u_i(t)$ are externally applied electromagnetic pulses that can be used to manipulate, or guide, the evolution of the system state. Typical pulse design problems involve designing these pulses, or controls, to bring the final state of the density matrix $\rho(T)$ as close as possible to a target operator. This problem can be transformed, by taking the expectation values of the operators involved in the state transfer, to the vector-valued, bilinear control problem, $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^m$ given by,

$$\frac{d}{dt}x = \left[\mathcal{H}_d + \sum_{i=1}^m u_i(t)\mathcal{H}_i \right] x, \quad (4.1)$$

where $\mathcal{H}_d \in \mathbb{R}^{n \times n}$ corresponds to the drift evolution representing H_f and L , $\mathcal{H}_i \in \mathbb{R}^{n \times n}$ corresponds to the controlled evolution representing H_i , and $t \in [0, T]$ [41]. While (4.1) accurately represents the classical interaction of magnetic fields, in practice the effective fields - and, therefore, the matrices representing the Hamiltonians \mathcal{H}_d and \mathcal{H}_i - show variation in magnitude due to different chemical environments and equipment errors. The system can no longer be described by a single equation but rather by a family of equations with variation in the parameters that characterize the motion, which motivates us to consider the dispersion in the dynamics as a continuum indexed by the system values,

$$\frac{d}{dt}x(t, s) = \left[\mathcal{H}_d(s) + \sum_{i=1}^m u_i(t) \mathcal{H}_i(s) \right] x(t, s), \quad (4.2)$$

where $s \in \mathcal{S} \subset \mathbb{R}^d$ is a d -dimensional interval representing the d parameters exhibiting variation [56]. In a more general formulation the matrices representing the Hamiltonians can be time-dependent, $\mathcal{H}_d = \mathcal{H}_d(t, s)$ and $\mathcal{H}_i = \mathcal{H}_i(t, s)$, as in the case of random fluctuations. Designing a single set of controls (pulses) $u_i(t)$ that simultaneously steer an ensemble of dispersive systems in (4.2) from an initial state to a desired final state is a fundamental problem in the control of quantum systems.

4.3 Pulse Design on the Bloch Equations

We first consider several examples based on the prototypical quantum control system described by the Bloch equations [3]. The Bloch equations have been found to model a range of quantum phenomena from protein spectroscopy in nuclear magnetic resonance (NMR) [15] and medical scans in magnetic resonance imaging (MRI) [13] to Rabi oscillations in quantum optics [55]. In the following discussion, we will consider the specific application and terminology for NMR spectroscopy in liquids, however, the methods and results are easily transferred to these other areas of interest. In NMR spectroscopy, when the duration of the pulse design problem is small compared with the relaxation times ($T \ll T_1, T_2$, the characteristic longitudinal and transverse relaxation times, respectively), the evolution of spins can be well approximated as sequences of unitary rotations driven by the static magnetic field and the applied

electromagnetic controls. In practice, the effective fields generating these rotations show variation across the quantum sample due to hardware imperfection and chemical shielding, which leads us to consider a range of magnetic field variations. The corresponding dimensionless Bloch equations in the rotating frame (see Appendix C) are,

$$\frac{d}{dt}M(t, \omega, \epsilon) = \left[\omega \Omega_z + \epsilon u(t) \Omega_y + \epsilon v(t) \Omega_x \right] M(t, \omega, \epsilon), \quad (4.3)$$

where $M(t, \omega, \epsilon) = (M_x(t, \omega, \epsilon), M_y(t, \omega, \epsilon), M_z(t, \omega, \epsilon))$ is the Cartesian magnetization vector for the parameter values $s = (\omega, \epsilon)$, $\omega \in [-B, B] \subset \mathbb{R}$, is the dispersion of natural frequencies, $\epsilon \in [1 - \delta, 1 + \delta]$, $0 < \delta < 1$, is the amplitude attenuation factor, and $\Omega_\alpha \in \text{SO}(3)$ is the generator of rotation around the α axis. A pulse that compensates for the dispersion in frequency and is insensitive to the scaling of the applied controls is called a broadband pulse robust to rf inhomogeneity. In this section we consider several examples based on this model, including pulses robust not only to frequency dispersion and inhomogeneity, but also robust to uncertainty in initial conditions and time-varying frequencies.

4.3.1 Broadband Excitation and Inversion

A canonical problem in the control of quantum systems modeled by the Bloch equations is to design pulses that will accomplish a state-to-state transfer of the system. Such pulses, e.g., $\pi/2$ and π pulses (accomplishing $\pi/2$ and π rotations, respectively), are the fundamental building blocks of the pulse sequences used in many quantum experiments. First consider the excitation, or $\pi/2$, pulse that rotates the net magnetization from the equilibrium position ($+z$) to the $+x$ axis, i.e., $M(0) = (0, 0, 1)' \rightarrow M(T) = (1, 0, 0)'$. In the ensemble case, this goal corresponds to a uniform excitation of the spin vector across all choices of frequency and inhomogeneity. Specifically we

consider the optimal ensemble control problem,

$$\begin{aligned}
\min \quad & \alpha_1 \int_{1-\delta}^{1+\delta} \int_{-B}^B -M_x(T, \omega, \epsilon) d\omega d\epsilon + \alpha_2 \int_0^T u^2(t) + v^2(t) dt, \\
\text{s.t.} \quad & \frac{d}{dt} M(t, \omega, \epsilon) = \left[\omega \Omega_z + \epsilon u(t) \Omega_y + \epsilon v(t) \Omega_x \right] M(t, \omega, \epsilon), \\
& M(0, \omega, \epsilon) = (0 \ 0 \ 1)', \\
& \sqrt{u^2(t) + v^2(t)} \leq A, \ \forall t \in [0, T],
\end{aligned} \tag{4.4}$$

where A is the maximum allowable amplitude and the cost functional serves to minimize the z-component of the spin vector (integrated across the ensemble) and the energy of the designed pulse, with relative weighting given by α_1 and α_2 (unless otherwise mentioned $\alpha_1 = 1$ and $\alpha_2 = 0$). A similar cost can be used for a π pulse designed to steer the magnetization to $(0, 0, -1)'$, by replacing $-M_x$ by M_z .

Figure 4.1 presents both a broadband $\pi/2$ and π pulse designed by the multidimensional pseudospectral method without consideration of rf inhomogeneity ($\epsilon = 1$). The corresponding $\pi/2$ and π resonance offset profiles, simulated and experimental, are plots of $M_x(T, \omega, 1)$ and $M_z(T, \omega, 1)$, respectively, over the frequency offset ω . The optimized band, here $\omega \in [-A, A]$, with maximum amplitude $A = 20$ kHz, in the experimental profile is marked by the overline. This thesis, to be both relevant to those from both chemical physics and control backgrounds, will interchangeably use physical units and normalized ones to refer to physical parameters such as amplitudes, bandwidths, and durations (see Appendix C for details). The dips present in the π pulse experimental profile indicate the presence of rf inhomogeneity as shown in the simulated profile for various choices of ϵ . Such discrepancies motivate the need for compensation not only in frequency but also in rf scaling.

It is important to compare the derived pulses with some of the conventional approaches used in NMR experiments. The resonance offset profile for a π pulse designed for a wider bandwidth $\omega \in [-2A, 2A]$, $A = 20$ kHz shows marked improvement over the conventional hard pulse in Figure 4.2. The hard pulse is a maximum amplitude on-resonance pulse with duration $T = \pi/2A$, which performs well for the center (on-resonance) frequency and poorer at the edges of the bandwidth.

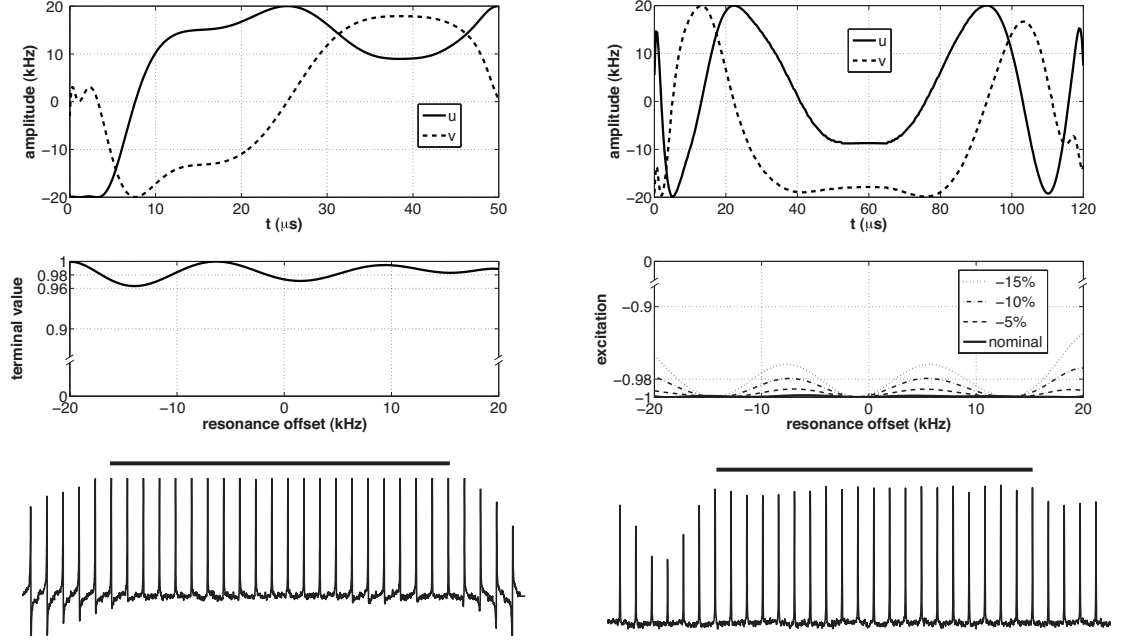


Figure 4.1: The control pulse shape (top), simulated resonance offset profile (middle), and experimental resonance offset profile (bottom) for an excitation (left) and inversion pulse (right) developed with the multidimensional pseudospectral method. The $\pi/2$ pulse achieves an average x component of 0.9852 with the parameters: $A = 20$ kHz, $B = 20$ kHz, $T = 50\mu\text{s}$, $N = 32$, $N_\omega = 8$. The π pulse achieves an average z component of -0.9991 with the parameters: $A = 20$ kHz, $B = 20$ kHz, $T = 120\mu\text{s}$, $N = 36$, $N_\omega = 12$. The difference between the simulated and experimental offset profiles highlights the effect of rf inhomogeneity (simulated response is shown for $\epsilon = 0.85, 0.9, 0.95, 1$).

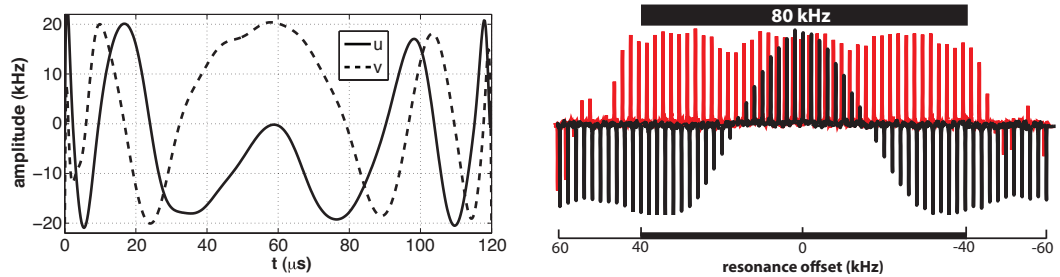


Figure 4.2: A broadband π pulse and the corresponding excitation profile (red) compared to the excitation profile of the conventional hard pulse (black). The dips in the excitation profile are due to rf inhomogeneity in the experimental equipment. $A = 20$ kHz, $B = 40$ kHz, $T = 120\mu\text{s}$, $N = 36$, $N_\omega = 12$.

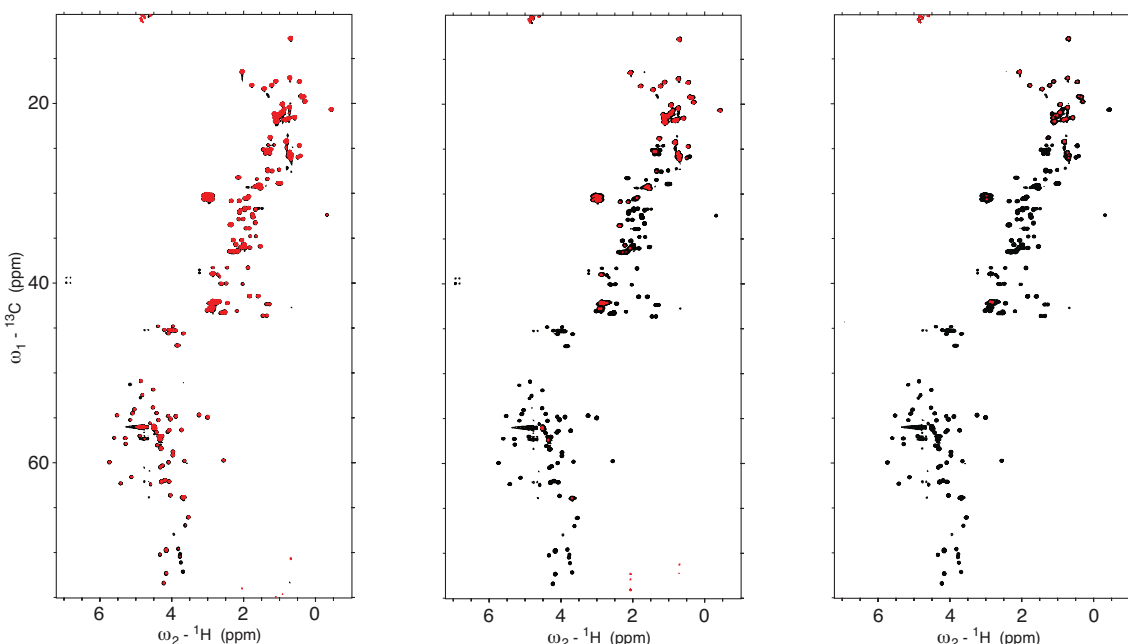


Figure 4.3: The 2D spectra from on-resonance (black) and 250 ppm off-resonance (red) HSQC experiments using the pulse in Figure 4.2 (left), a standard adiabatic pulse (middle), and the hard pulse (right). The enhanced signal off-resonance enables better determination of protein structure and the optimized pulse achieves up to 20 times the sensitivity of the adiabatic pulse and with shorter duration ($120\mu\text{s}$ versus $500\mu\text{s}$).

If the pulse in Figure 4.2 is incorporated into an HSQC experiment, the resulting 2D spectra shows enhanced signal recovery off-resonance of up to 20 times compared to the same experiment using an adiabatic pulse and higher still for the hard pulse [42]. Figure 4.3 presents the on-resonance (black) and off-resonance (red) 2D spectra for these pulse choices. The on-resonance spectra is expected to be the same, as each pulse works well at the nominal (resonance) frequency. However, the reduced off-resonance spectra is caused by poor broadband inversion at the edges of the bandwidth. Such plots are used to determine protein structure by investigating the chemical shifts present in the molecule. It is crucial to reduce signal loss so that an accurate assessment can be made of the experiment and a reliable model of the protein can be constructed.

One of the hallmark features of the pseudospectral method is the freedom and flexibility to specify any cost function and any constraint. Several computational methods rely on ad hoc techniques at each step in the optimization to project the state and controls back onto the admissible set of states and controls [34]. For example, to yield a pulse of desired smoothness with the Krotov method, it is necessary to perform a Fourier truncation after each step in the optimization in order to eliminate high frequency components [45]. The collocation involved in the pseudospectral approximation imposes the constraints in a very straightforward fashion and serve as bounds on the exploration of the state space, so no ad hoc methods are necessary.

The π pulse in Figure 4.4 was optimized by the method with an imposed symmetry constraint [58]. Such a condition, which is often observed in π pulses calculated without the constraint, reduces the number of control variables in the optimization by equating $\bar{u}_k = \bar{u}_{N-k}$ and likewise for v . In addition, the pulse in Figure 4.4 was designed allowing the terminal time to be a variable of the optimization. Other methods require many runs, testing different pulse durations, to determine the optimal time, whereas here this can be achieved in a single optimization.

In these past examples, we have considered broadband pulse designs only and have seen the effect of rf inhomogeneity in the experimental offset profiles. To compensate for the rf scaling inherent in the experiments, we consider (4.4) with a two dimensional variation in both ω and ϵ . Figure 4.5 compares two such pulses derived by the multidimensional pseudospectral method under different cost functions. The standard choice ($\alpha_1 = 1$ and $\alpha_2 = 0$) achieves an average excitation of 0.98. The minimum energy pulse ($\alpha_1 = 0$ and $\alpha_2 = 1$) is optimized with an additional constraint requiring the average excitation to be greater than the previous optimization, i.e., $\sum_{i,j} M_x(T, \omega_i, \epsilon_j) / N_\omega N_\epsilon \geq 0.98$. This pulse achieves an equivalent performance

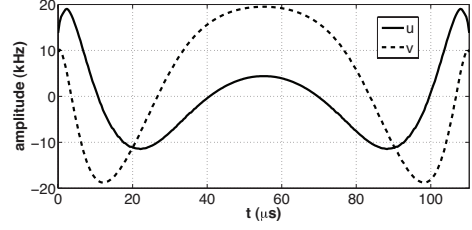


Figure 4.4: Arbitrary constraints are easy to include in the pseudospectral formational for pulse design, such as with a symmetry constraint on the control or leaving the pulse duration free to vary, i.e. $0 < T \leq T_{\max}$. The average z magnetization of the offset (not shown) is -0.9923 . Parameters: $A = 20$ kHz, $B = 20$ kHz, $T_{\max} = 200\mu\text{s}$, $N = 24$, $N_\omega = 9$.

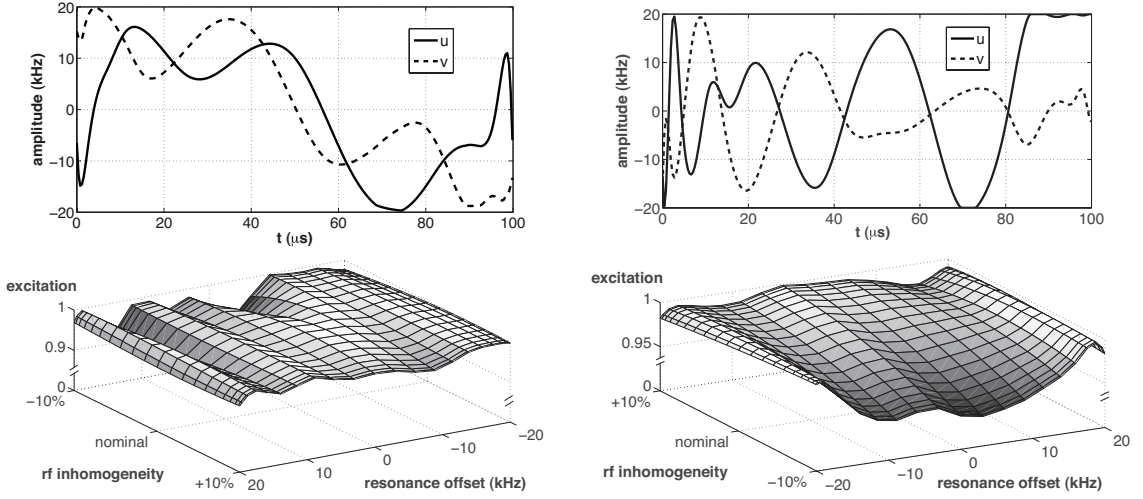


Figure 4.5: Broadband excitation $\pi/2$ pulses designed to compensate for 10% rf inhomogeneity with average excitation 0.98. The minimum-energy broadband pulse (right) is optimized by solely minimizing energy subject to a performance constraint and achieves this transfer with 16% less rf energy. Parameters: $A = 20$ kHz, $B = 20$ kHz, $\delta = 0.1$, $T = 100\mu\text{s}$, $N = 24$, $N_\omega = 8$, $N_\epsilon = 1$.

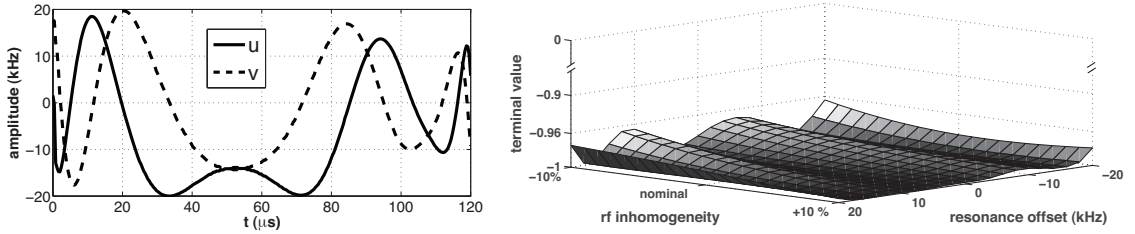


Figure 4.6: Broadband inversion π pulses designed to compensate for 10% rf inhomogeneity with average excitation -0.9929 . The pulse was optimized to minimize both magnetization and rf energy. Parameters: $A = 20$ kHz, $B = 20$ kHz, $\delta = 0.1$, $T = 120\mu\text{s}$, $N = 36$, $N_\omega = 8$, $N_\epsilon = 1$.

(within 0.3%) with 16% less rf energy, as given by the second term in the cost of (4.4). Similarly, an inversion pulse robust to variation in frequency and rf scaling can be optimized as in Figure 4.6 with equal weight given to minimizing average inversion and control energy ($\alpha_1 = 1$ and $\alpha_2 = 1$).

4.3.2 Variation in Initial Conditions

In most experiments, individual pulses, such as the ones in the preceding section, are combined into a longer pulse sequence, which performs a more complicated manipulation of the system state with intermediate steps and goals. Even in the case of highly optimized individual pulses, as shown previously, there is an error between the desired and actual final states. Moreover, pulses depend upon an exact (and usually uniform) initial condition in order to achieve their expected levels of performance. These effects combine to create a magnified accumulated error at the termination of the pulse sequence. The variation of the initial conditions of these pulses, therefore, causes significant degradation in achievable performance.

A representative example of such a pulse sequence is to perform a three step pulse sequence, which rotates the magnetization of the ensemble (1) from equilibrium ($+z$) to a point on the transverse plane (e.g. $+y$); (2) to the opposite point on the transverse plane (e.g. $-y$); (3) back to the equilibrium position ($+z$). Such pulses generally include “phase locking” pulses before and after the second pulse during which the magnetization dissipates. This dissipation is the portion of the experiment that is important to recover accurately and reflects a quantity to be measured, for example, a metabolic rate [48, 51]. If, in addition, there is accumulated error due to uncertainty in the initial conditions of the individual pulses, this leads directly to measurement inaccuracy. Here, by removing the “phase locking” pulses, we can abstract this pulse sequence to a unitary process and directly address any losses due to error. The controllability of the Bloch equations is shown by constructing parameter-dependent (e.g. frequency, rf inhomogeneity) rotations of the spin vectors [38]. This, therefore, ensures that the problem with variation in initial conditions can be solved provided that the initial conditions can be parameterized by the frequency and rf inhomogeneity.

Figure 4.7 displays a three-stage optimized pulse designed by the multidimensional pseudospectral method which is robust to frequency dispersion and variation in the initial conditions of the three stages. This pulse was run as three concurrent optimizations, with the final states of one pulse fed in as the initial conditions of the next. This optimized pulse is compared with the combination of three separately optimized pulses; these combined pulses were designed with equal total duration. The terminal profiles at each intermediate goal quickly show the evidence of accumulated error in

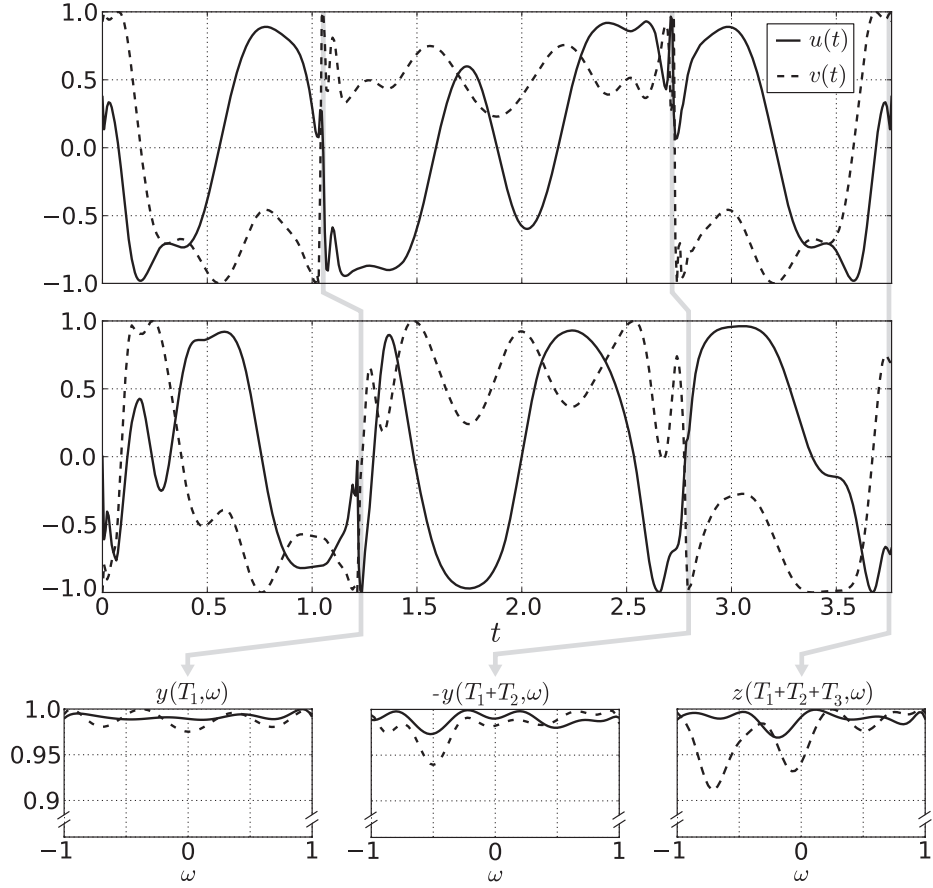


Figure 4.7: Pulses are optimized to produce a desired $z \rightarrow y \rightarrow -y \rightarrow z$ evolution of the Bloch equations. The upper plot displays the concatenation of individually optimized $z \rightarrow y$ and $y \rightarrow -y$ pulses, which achieves the dashed terminal profiles shown below, with respective average performances: 0.99, 0.98, 0.97 (0.91 minimum). The middle plot displays a 3-part simultaneously-optimized pulse robust to variation in the initial condition and achieves the solid terminal profiles shown below, with respective average performances: 0.99, 0.99, 0.99 (0.97 minimum). The noticeable enhancement in performance and uniformity is due to compensating for the inhomogeneity in the initial condition of the individual pulses.

the case of the individually optimized pulses (each individual pulse has an average performance greater than 0.98). Most importantly, the uniformity of the inversion is lost in the additive error, with dips in performance down to 0.91.

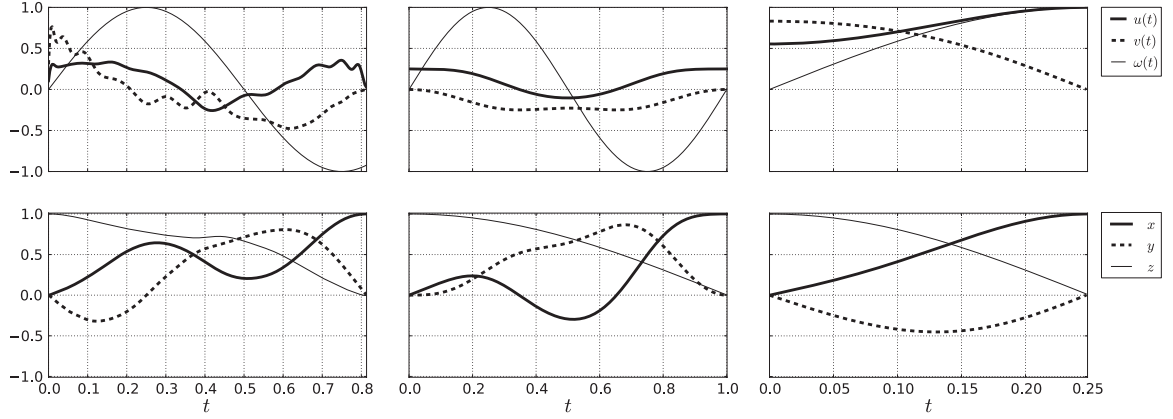


Figure 4.8: Control pulses (top) and state trajectories (bottom) corresponding to different objectives and designed to compensate for the time-varying frequency $\omega(t) = \sin(t)$. A single-system state transfer $M(0) = (0, 0, 1)' \rightarrow M(T) = (1, 0, 0)'$ is designed using the terminal cost $\varphi(T) = M_x(T)$ and running costs $\mathcal{L}(t) = 0$ (left), $\mathcal{L}(t) = 0.1(u(t)^2 + v(t)^2)$ (middle), $\mathcal{L}(t) = 0.1$ (right). The terminal time was free in all cases, bounded by $T_{\max} = 1$.

4.3.3 Time-Varying Frequency

Until now, we have considered that the dispersion and uncertainty of the system are stationary. However, addressing time-varying fluctuations in parameters is also of particular theoretical and practical importance. For example, in the formulation of quantum control problems given in (4.2) we noted that the Hamiltonians can be time-varying, motivated by such phenomena as random telegraph noise [49]. The first step to addressing stochastic variations in such physical systems is to demonstrate control of time-varying systems, such as given by the expectation value of the corresponding random process.

Figure 4.8 presents a series of optimizations designing $\pi/2$ pulses providing a state transfer $+z$ to $+x$, while compensating for a time-varying frequency, $\omega(t) = \sin(t)$. Various choices of cost functional yield different results. The arbitrary control pulse profile corresponding to the terminal cost $\varphi(T) = M_x(T)$ (Fig. 4.8, left) motivates studying optimal control methods that provide the capacity for hybrid objectives resulting in more physically meaningful controls, e.g. minimizing energy (middle) and time (right).

4.4 Relaxation Optimized Pulse Design

Multidimensional NMR experiments for protein spectroscopy rely on the concept of polarization transfer, in which coherence is moved from one spin to another through the application of appropriate pulse sequences. Two dimensional spectra, as shown in Figure 4.3, are generated by observing the transfer of polarization from one spin (e.g., hydrogen, spin I_1) to another (e.g., carbon, spin I_2). Initially, spins I_1 are excited by an applied pulse that rotates the magnetization to the transverse plane. Then, following a delay t_1 , a polarization transfer pulse is applied, in which magnetization is moved from spins I_1 to I_2 via the coupling between the spins. Spins I_2 are then likewise excited and their precession is recorded after another delay t_2 . Using the delay values and the recorded signal, the efficiency of transfer can be plotted, corresponding to the pair $(\omega_{I_1}, \omega_{I_2})$, which creates the contour plot of a 2D spectra.

We study two important variants of this problem as representative examples of systems in which maximum efficiency is sought in the presence of dissipation. Relaxation effects could be neglected in the Bloch equations due to the short timescales of the developed pulses. In this application, however, relaxation is a critical component that limits the signal recovery in real experiments. Naively, with a goal to maximize a terminal state that is subject to dissipation, it seems appropriate to cast the problem as a minimum time optimization. However, deeper investigation reveals that these are not necessarily optimal and that an interplay between the states of the system can yield increased signal. This observation underscores the study of relaxation optimized pulse design.

The polarization transfer problem from liquid state NMR is composed of a pair of isolated, heteronuclear $1/2$ spin particles, which we label I_1 (for example ^1H) and I_2 (for example ^{13}C or ^{15}N), with a scalar coupling J that quantifies the interaction between the spins [20]. In a doubly rotating frame, which rotates with each spin at its resonance frequency, the free evolution Hamiltonian for this system is $H_f = 2JI_{1z}I_{2z}$, where $I_{1z} = \sigma_{1z}/2$, $I_{2z} = \sigma_{2z}/2$ and σ_{1z}, σ_{2z} are the Pauli spin matrices for spins I_1 and I_2 respectively. This Hamiltonian assumes that $|\omega_1 - \omega_2| \gg J$ (weak coupling limit relative to the resonance frequencies of the spins) so that the Heisenberg coupling ($\mathbf{I}_1 \cdot \mathbf{I}_2$) can be well approximated by the scalar coupling ($I_{1z}I_{2z}$) [20].

As we have seen with the Bloch equations, the relaxation rates and scalar coupling depend upon the physical constants of the spins, such as gyromagnetic ratios, inter-nuclear distance, and the correlation time of the rotational tumbling [20]. However, these constants are scaled by the amplitudes of the various fields present in the system. Therefore, the same equipment and location inhomogeneities cause variation in the values of the relaxation rates and coupling. It is then crucial to incorporate such variations into the models and optimizations for pulse design. We present the analysis and results for both the single and ensemble cases.

The most important relaxation mechanisms in NMR spectroscopy in liquid solutions are due to Dipole-Dipole (DD) interaction and Chemical Shift Anisotropy (CSA). We initially ignore the cross correlation caused by interference between DD and CSA relaxation.

4.4.1 Polarization Transfer without Cross-Correlated Relaxation

Here we consider the polarization transfer system with only DD and CSA relaxation ignoring the cross-correlated relaxation and study the case of slowly tumbling molecules in the spin diffusion limit, in which longitudinal relaxation rates ($1/T_1$) are negligible compared to transverse relaxation rates ($1/T_2$) [15].

The free evolution of the density matrix ρ representing the state of the coupled spin system in the doubly rotating frame is given by the master equation [32]

$$\begin{aligned} \dot{\rho} = & -iJ[2I_{1z}I_{2z}, \rho] - k_{DD}[2I_{1z}I_{2z}, [2I_{1z}I_{2z}, \rho]] \\ & - k_{CSA}^1[I_{1z}, [I_{1z}, \rho]] - k_{CSA}^2[I_{2z}, [I_{2z}, \rho]], \end{aligned} \quad (4.5)$$

where J is the scalar coupling constant, k_{DD} is the DD relaxation rate, and k_{CSA}^1, k_{CSA}^2 are CSA relaxation rates for spins I_1, I_2 , respectively.

The typical problem of interest involves designing pulses, ω_x and ω_y in the x and y axes respectively, to transfer the polarization from one spin to the other, i.e., $I_{1z} \rightarrow I_{2z}$. To do so, the polarization is typically passed through an intermediate, symmetric

step, $I_{1z} \rightarrow 2I_{1z}I_{2z} \rightarrow I_{2z}$. Because these two halves are symmetric, it is sufficient to consider only the first step of the transfer. In particular, the controls should be designed to maximize the final expectation value of the target operator $O = 2I_{1z}I_{2z}$ starting from $\rho(0) = I_{1z}$, i.e., $\langle 2I_{1z}I_{2z} \rangle(T) = \text{trace}\{\rho(T)2I_{1z}I_{2z}\}$, where T is the duration of the pulse and $\rho(T)$ is guided by the applied controls. We can recast this problem in matrix form [41],

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & -u_1 & 0 & 0 \\ u_1 & -\xi & -1 & 0 \\ 0 & 1 & -\xi & -u_2 \\ 0 & 0 & u_2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \quad (4.6)$$

where $x_1 = \langle I_{1z} \rangle$, $x_2 = \langle I_{1x} \rangle$, $x_3 = \langle 2I_{1y}I_{2z} \rangle$, $x_4 = \langle 2I_{1z}I_{2z} \rangle$ are expectations of the operators involved in the transfer; $\xi = (k_{DD} + k_{CSA}^1)/J$; and the controls $u_1(t) = \omega_y(t)/J$ and $u_2(t) = \omega_x(t)/J$ are the transverse components of the applied magnetic field, all normalized by the scalar coupling J .

We can now formulate the desired transfer objective as an optimal control problem on the dynamics given by (4.6). With polarization initially $x(0) = (1, 0, 0, 0)'$ our aim is to maximize $x_4(T)$ through the appropriate choice of $u_1(t)$ and $u_2(t)$. This problem has been solved analytically and the derived ROPE pulse establishes the maximum achievable transfer efficiency given by,

$$\eta_1 = \sqrt{\xi^2 + 1} - \xi. \quad (4.7)$$

Although the optimal control problem (4.6) does not consider the ensemble variations that occur in practice, applying the pseudospectral method reveals several interesting results. Figure 4.9 shows the transfer efficiency of several choices of ξ on the interval $[0, 1]$ compared with the optimal efficiency determined by ROPE. The pseudospectral method finds controls that have corresponding efficiencies that are negligibly close to the maximum efficiency. The figure also presents one such pulse solution and the corresponding trajectories, again comparing to the ROPE pulse.

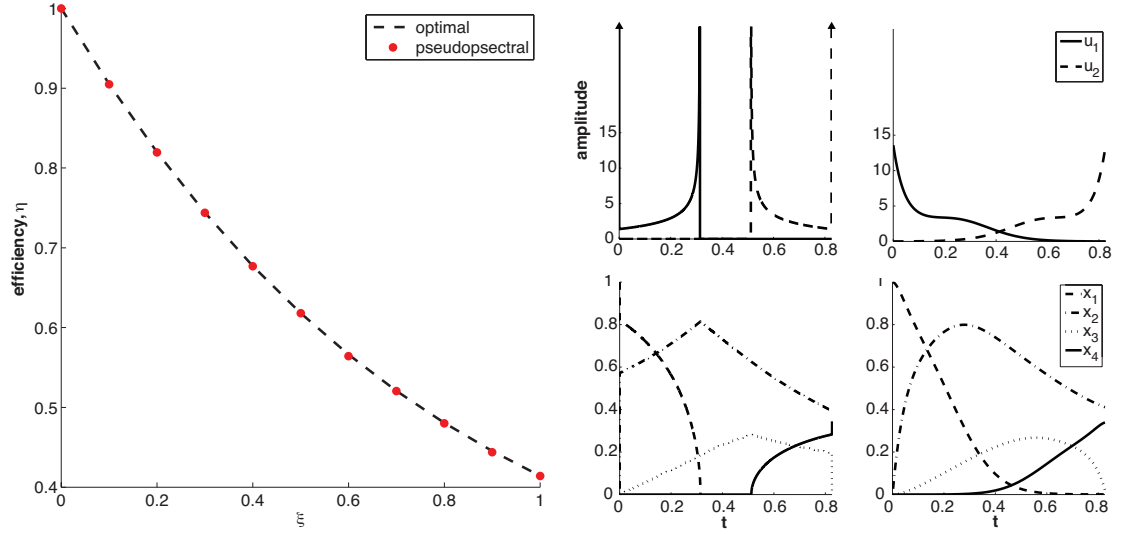


Figure 4.9: (Left) The efficiency of the transfer $x_1 \rightarrow x_4$ in system (4.6) achieved by the pseudospectral method, as a function of the relaxation parameter ξ in the range $[0, 1]$. The theoretically calculated maximum efficiency given by (4.7) is also shown. (Right) ROPE and Pseudospectral controls (top) and corresponding state trajectories (bottom) for $\xi = 1$. Each of the hard pulses within the ROPE pulse at $t=0$ and $t=T$ (top left) correspond to a 35° rotation and transfer the state from $x(0^-) = [1, 0, 0, 0]^T$ to $x(0^+) = [\cos 35^\circ, \sin 35^\circ, 0, 0]^T$ and from $x(T^-) = [0, x_2(T), \eta \sin 35^\circ, \eta \cos 35^\circ]^T$ to $x(T^+) = [0, x_2(T), 0, \eta]^T$, respectively, in near instantaneous time.

A clear advantage of the pseudospectral method well illustrated by this example is that the calculated control pulses are smooth functions. In contrast, notice the high-amplitude spikes at the beginning and end of each component of the analytic ROPE pulse. Such discontinuities can be challenging, if not impossible, to implement in practice and high amplitudes can be hazardous for the experiment sample, equipment, and human subjects (as in MRI). The pulse amplitude derived by the pseudospectral method, is easily implementable and maintains low values despite achieving transfer efficiencies within 1×10^{-3} of the theoretical optimal values. The pulse is attained from an optimization that minimizes energy subject to a constraint maintaining a desired transfer efficiency. Therefore, not only is the pseudospectral pulse without discontinuities but it also accomplishes the transfer with 45% less energy than the ROPE pulse.

We now turn our attention to the ensemble case, in which the system is modeled with variation in the relaxation rate and scalar coupling. We formulate the optimal ensemble control problem as

$$\begin{aligned}
\max \quad & \mathcal{J}_{\text{avg}} = \frac{1}{2\delta(\xi_2 - \xi_1)} \int_{1-\delta}^{1+\delta} \int_{\xi_1}^{\xi_2} x_4(T, \xi, J) d\xi dJ \\
\text{s.t.} \quad & \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & -u_1 & 0 & 0 \\ u_1 & -\xi & -J & 0 \\ 0 & J & -\xi & -u_2 \\ 0 & 0 & u_2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}, \\
& x(0) = [1, 0, 0, 0]', \\
& \sqrt{u_1^2(t) + u_2^2(t)} \leq A, \quad \forall t \in [0, T],
\end{aligned} \tag{4.8}$$

where \mathcal{J}_{avg} is the objective that maximizes the average final value of x_4 across the ensemble; $x_i = x_i(t, \xi, J)$ are again expectation values of the spin operators [57, 56]; T is the final time, free to vary as a decision variable; $\xi \in [\xi_1, \xi_2]$ is the transverse autocorrelated relaxation rate; $J \in [1 - \delta, 1 + \delta]$, $\delta \in [0, 1]$, is the scalar coupling constant; u_1 and u_2 are the applied controls; and A is the maximum allowable amplitude. In this case, these values have been normalized by the nominal scalar coupling J_0 .

Figure 4.10 illustrates smooth pseudospectral solutions to the normalized coherence transfer problem in (4.8) with $\xi \in [0, 2]$, fixed $J = 1$ (no variation in spin coupling), and amplitude bound $A = 20$. If we use the pseudospectral method with cost \mathcal{J}_{avg} we obtain the fluctuating pulse shown in Figure 4.10, top. If we use the same method with the hybrid cost $\mathcal{J} = \mathcal{J}_{\text{avg}} - \mathcal{J}_{\text{E}}$,

$$\mathcal{J}_{\text{E}} = \frac{1}{A^2 T} \int_0^T u_1^2(t) + u_2^2(t) dt.$$

we obtain the pulse shown in Figure 4.10, bottom, with less oscillation. Both of these pulses achieve a similar ensemble performance and the same figure depicts the coherence transfer corresponding to the hybrid objective pulse. Including the minimum energy term in the objective yields a significantly more implementable and physically

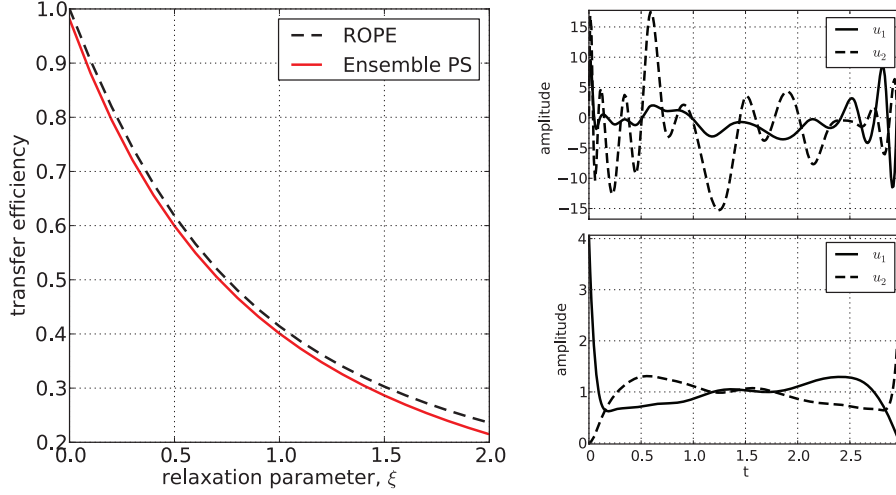


Figure 4.10: The optimal ensemble pulses (right) effectively compensate for all variations of ξ on the interval $[0, 2]$ with only minor losses in transfer efficiency (left) when compared to each analytic ROPE [32] efficiency for a single value of ξ ($N = 28$ and $N_\xi = 8$). The top optimal ensemble pulse was developed by maximizing the average transfer efficiency and the pulse beneath was developed by maximizing the average transfer efficiency and minimizing energy. The transfer efficiency plot corresponds to the latter.

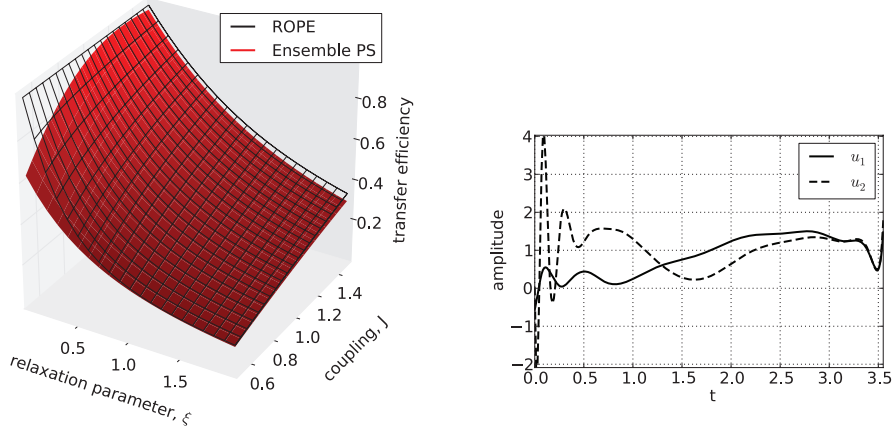


Figure 4.11: The optimal ensemble pulse shown in (right) effectively compensates for all variations of $\xi \in [0, 2]$ and $J \in [0.5, 1.5]$ with comparable efficiency (left) to each ROPE pulse [32] for a specific ξ and J . This optimal ensemble pulse was developed by maximizing average transfer efficiency and minimizing energy with $N = 24$, $N_\xi = 8$, and $N_J = 4$.

intuitive pulse. In most cases there are a large (possibly uncountable) number of feasible solutions that achieve a similar performance and it is experimentally advantageous to select from this large number the one that also minimizes energy.

Figure 4.11 shows a solution to the problem posed in (4.8) and corresponding coherence transfer for the two-dimensional ensemble problem with $\xi \in [0, 2]$ and $J \in [0.5, 1.5]$ using the same hybrid objective defined above. The transfer of the ensemble pulse derived with the pseudospectral method is robust to both parameter variations and still compares favorably with the upper bound achieved by the ROPE pulses.

4.4.2 Polarization Transfer with Cross-Correlated Relaxation

If DD-CSA cross-correlated relaxation cannot be neglected, the master equation as in (4.5) is then modified to incorporate it as [30]

$$\begin{aligned}\dot{\rho} = & -iJ[2I_{1z}I_{2z}, \rho] - k_{DD}[2I_{1z}I_{2z}, [2I_{1z}I_{2z}, \rho]] \\ & - k_{CSA}^1[I_{1z}, [I_{1z}, \rho]] - k_{CSA}^2[I_{2z}, [I_{2z}, \rho]] \\ & - k_{DD/CSA}^1[2I_{1z}I_{2z}, [I_{1z}, \rho]] - k_{DD/CSA}^2[2I_{1z}I_{2z}, [I_{2z}, \rho]],\end{aligned}$$

where k_{DD} , k_{CSA}^1 , k_{CSA}^2 are auto-relaxation rates due to DD relaxation, CSA relaxation of spin I_1 , CSA relaxation of spin I_2 and $k_{DD/CSA}^1$, $k_{DD/CSA}^2$ are cross-correlation rates of spins I_1 and I_2 due to interference between DD and CSA relaxation mechanisms.

Using this master equation, we can find the corresponding matrix equation, which requires a larger state space to represent it. Here we jump directly to the ensemble case where we consider variation in the autocorrelated relaxation rate,

$$\begin{aligned}\max \quad & \mathcal{J}_{\text{avg}} = \frac{1}{(\xi_2 - \xi_1)} \int_{\xi_1}^{\xi_2} x_6(T, \xi_a) d\xi_a \\ \text{s.t.} \quad & \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \\ \dot{x}_5 \\ \dot{x}_6 \end{bmatrix} = \begin{bmatrix} 0 & -u_1 & u_2 & 0 & 0 & 0 \\ u_1 & -\xi_a & -0 & -1 & -\xi_c & 0 \\ -u_2 & 0 & -\xi_a & -\xi_c & 1 & 0 \\ 0 & 1 & -\xi_c & -\xi_a & 0 & -u_2 \\ 0 & -\xi_c & -1 & -0 & -\xi_a & u_1 \\ 0 & 0 & 0 & u_2 & -u_1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}, \\ & x(0) = [1, 0, 0, 0, 0, 0]', \\ & \sqrt{u_1^2(t) + u_2^2(t)} \leq A, \quad \forall t \in [0, T],\end{aligned}\tag{4.9}$$

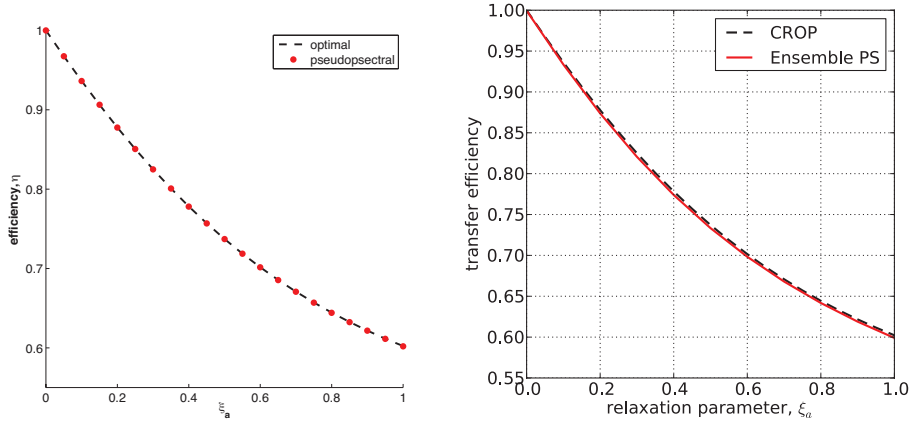


Figure 4.12: The efficiency (left) of the transfer $x_1 \rightarrow x_6$ in system (4.9) without variation achieved by the pseudospectral method, as a function of the relaxation parameter ξ_a in the range $[0, 1]$, with $\xi_c = 0.75\xi_a$. The efficiency (right) of the ensemble case closely reproduces the optimal efficiency with a single pulse for the entire $[0, 1]$ range.

where $x_1 = \langle I_{1z} \rangle$, $x_2 = \langle I_{1x} \rangle$, $x_3 = \langle I_{1y} \rangle$, $x_4 = \langle 2I_{1y}I_{2z} \rangle$, $x_5 = \langle 2I_{1x}I_{2z} \rangle$, $x_6 = \langle 2I_{1z}I_{2z} \rangle$; $\xi_a = (k_{DD} + k_{CSA}^1)/J \in [\xi_1, \xi_2]$; $\xi_c = k_{DD/CSA}^1/J$; and $u_1(t), u_2(t)$ are the available controls as before [56]. Starting from $x(0) = (1, 0, 0, 0, 0, 0)'$, we want to design $u_1(t)$ and $u_2(t)$ that maximize $x_6(T)$ subject to the dynamics given above.

This problem has also been solved analytically and the analytical pulse was denoted as CROP [30], which attains the maximum achievable value of x_6 , i.e., the efficiency η_2 of the transfer is given by the same form as before

$$\eta_2 = \sqrt{\xi^2 + 1} - \xi, \quad (4.10)$$

but now

$$\xi = \sqrt{\frac{\xi_a^2 - \xi_c^2}{1 + \xi_c^2}}. \quad (4.11)$$

Similar to the previous problem, the pseudospectral and multidimensional pseudospectral method performs well, matching the analytic efficiency in both cases, see Figure 4.12. The optimization considering the ensemble variation achieves a level of polarization transfer with an average of 3×10^{-3} deviation from the optimal upper bound with only $N = 16$ and $N_{\xi_a} = 2$. Maintaining low sampling numbers is of

particular interest as the computational complexity of an ensemble problem of high dimension grows quickly.

Chapter 5

Convergence of the Pseudospectral Method

By accepting and implementing a numerical method, we implicitly assume that the transformations and discretization used to prepare the problem for computational work does not fundamentally alter the nature of the problem. It is then critically important to show that this assumption is justified. Here we do so by both empirical and theoretical means. More specifically, we show that as the number of discretizations in the pseudospectral method (and samples in the multidimensional pseudospectral method) increases the solution of the algebraic nonlinear programming problem converges to the solution of the original continuous-time optimal control problem. For this argument, we consider a modified nonlinear programming problem statement.

Without loss of generality, we consider a general continuous-time optimal control problem defined on the time interval $\Omega = [-1, 1]$, which can be achieved by a simple affine transformation.

Problem 1 (Continuous-Time Optimal Control).

$$\min \quad J(x, u) = \varphi(x(1)) + \int_{-1}^1 \mathcal{L}(x(t), u(t)) \, dt, \quad (5.1)$$

$$\text{s.t.} \quad \frac{d}{dt}x(t) = f(t, x(t), u(t)), \quad (5.2)$$

$$e(x(-1), x(1)) = 0, \quad (5.3)$$

$$g(x(t), u(t)) \leq 0, \quad (5.4)$$

$$\|u(t)\|_{\infty} \leq A, \quad u \in H_m^{\alpha}(\Omega), \quad \alpha > 2 \quad (5.5)$$

where $\varphi \in C^0$ is the terminal cost; the running cost, $\mathcal{L} \in C^\alpha$, where C^α is the space of continuous functions with α classical derivatives, and dynamics, $f \in C_n^{\alpha-1}$, where $C_n^{\alpha-1}$ is the space of n -vector valued $C^{\alpha-1}$ functions, with respect to the state, $x(t) \in \mathbb{R}^n$, and control, $u(t) \in \mathbb{R}^m$; e and g are terminal and path constraints, respectively; $H_m^\alpha(\Omega)$ is the m -vector valued Sobolev space. The norm associated with the Sobolev space with $m = 1$, $H^\alpha(\Omega)$, is given with respect to the $L^2(\Omega)$ norm [10],

$$\|h\|_{(\alpha)} = \left(\sum_{k=0}^{\alpha} \|h^{(k)}\|_2^2 \right)^{1/2}.$$

Problem 2 (Algebraic Nonlinear Programming).

$$\min \quad \bar{J}(\bar{x}, \bar{u}) = \varphi(\bar{x}_N) + \sum_{k=0}^N \mathcal{L}(\bar{x}_k, \bar{u}_k) w_k \quad (5.6)$$

$$\text{s.t.} \quad \|f(\mathcal{I}_N x, \mathcal{I}_N u) - \mathcal{D}_N x\|_N \leq c_d N^{1-\alpha} \quad (5.7)$$

$$e(\bar{x}_0, \bar{x}_N) = 0 \quad (5.8)$$

$$g(\bar{x}_k, \bar{u}_k) \leq 0 \quad (5.9)$$

$$\|u_k\| \leq A \quad \forall k = 0, 1, \dots, N \quad (5.10)$$

where c_d is a positive constant; we define the discrete $L_n^2(\Omega)$ norm $\|h\|_N = \sqrt{\langle h, h \rangle_N}$, for $h, h_1, h_2 \in L_n^2(\Omega)$, $\Omega = [-1, 1]$, with,

$$\langle h_1, h_2 \rangle_N = \sum_{k=0}^N h_1'(t_k) h_2(t_k) w_k,$$

where w_k is the Gauss quadrature weight from (3.1).

Remark 1. The dynamics in (5.7) have been relaxed from the equality in (3.9) to ensure the feasibility of the discrete problem, which is used in Proposition 1. It is trivial to show that in the limit, as $N \rightarrow \infty$, these two conditions coincide.

We seek to address three questions related to solving the continuous-time optimal control (Problem 1) by solving the pseudospectral discretized constrained optimization (Problem 2). Suppose a feasible solution (x, u) exists to Problem 1. Under what conditions:

1. **Feasibility:** For a given order of approximation, N , does Problem 2 have a feasible solution, (\bar{x}, \bar{u}) , which are the interpolation coefficients given in (3.4) and (3.5)?
2. **Convergence:** As N increases, does the sequence of optimal solutions, $\{(\bar{x}^\dagger, \bar{u}^\dagger)\}$, to Problem 2 have a corresponding sequence of interpolating polynomials which converges to a feasible solution of Problem 1? Namely,

$$\lim_{N \rightarrow \infty} (\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger) = (x, u)$$

3. **Consistency:** As N increases, does the convergent sequence of interpolating polynomials corresponding to the optimal solutions of Problem 2 converge to an optimal solution of Problem 1? Namely,

$$\lim_{N \rightarrow \infty} (\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger) = (x^*, u^*)$$

Remark 2. *It is possible that Problem 1 has more than one optimal solution, i.e., there is more than one solution with the same optimal cost $J(x^*, u^*) = J^*$. Therefore, to show that the sequence of discrete solutions converges to an optimal solution, we can instead show that the cost of the discrete solution, \bar{J} , converges to the optimal cost J^* .*

Previous work has been done in the area of convergence of the pseudospectral method and we aim to augment this literature with several key insights that make convergence results applicable to a wider class of systems and relax the conditions on which the current proofs are based. Rather complete analysis has been done for the class of nonlinear systems which can be feedback linearized, including convergence rates [21]. We show below that ensemble quantum systems of interest do not fall within the class of feedback linearizable systems. Work has also included general nonlinear systems, but with the assumption that the solutions of the algebraic nonlinear programming problem have a limit point (i.e., have a convergent subsequence) [22]. In the language used above, this is very close to assuming “Convergence”, which in this presentation we relax and prove Feasibility, Convergence, and Consistency directly. Finally, we examine the convergence of the multidimensional pseudospectral method as applied

to ensemble optimal control problems. In what follows we consider first the convergence of the standard pseudospectral method and then discuss the convergence of the ensemble case.

We first observe that ensemble control systems of interest are not feedback linearizable [26], which motivates a need for a more general convergence proof. Consider the bilinear Bloch equations in (4.3) without variation in rf inhomogeneity (i.e., $\epsilon = 1$). The ability to feedback linearize a general nonlinear system is given by the Lie algebra generated by the drift and control vector fields (the conditions on this algebra must hold for each control term individually; here we consider the case for u). In particular, the terms $\text{ad}_{\omega\Omega_z}^0 \Omega_y = \Omega_y$, $\text{ad}_{\omega\Omega_z}^1 \Omega_y = -\omega\Omega_x$, $\text{ad}_{\omega\Omega_z}^2 \Omega_y = -\omega^2\Omega_y$, \dots , and,

$$\begin{aligned}\text{ad}_{\omega\Omega_z}^{2k-1} \Omega_y &= (-1)^k \omega^{2k-1} \Omega_x, \\ \text{ad}_{\omega\Omega_z}^{2k} \Omega_y &= (-1)^k \omega^{2k} \Omega_y,\end{aligned}$$

where $k = 1, 2, \dots$, and ω is any value in the interval $\mathcal{S} \subset \mathbb{R}$. It is clear that this Lie algebra, with increasing powers of the parameter ω , is never closed. Therefore, the span of the appropriate Lie brackets is not involutive, which indicates that such a system is not feedback linearizable.

5.1 Empirical Convergence

Here we present convergence results for the pseudospectral optimizations presented in this paper. The convergence property is related to the conditions under which a sequence of discretized optimization solutions, provided existing, converges to the original optimal control solution as the number of nodes (discretizations) increases. Since analytic convergence results are challenging to identify for general systems, we first analyze the convergence numerically to guarantee that solutions do converge as the number of nodes, N , increases.

The orthogonal polynomials of the pseudospectral method provide spectral convergence rates similar to Fourier series approximations for periodic functions, which can easily be seen in practice. Figure 5.1 demonstrates the characteristic convergence

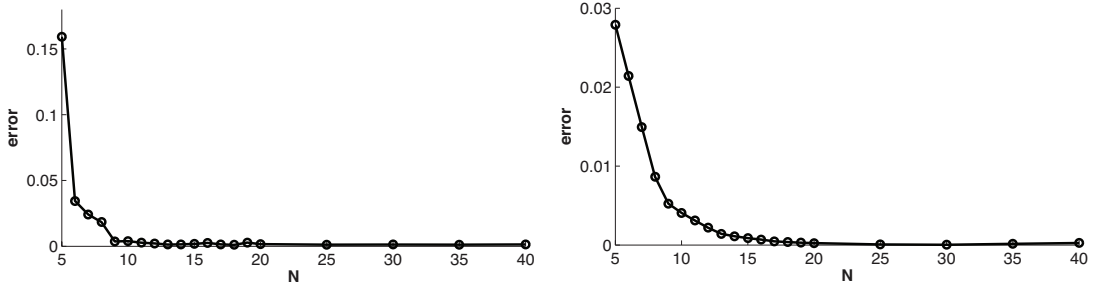


Figure 5.1: Numerical results are shown for the convergence of the pseudospectral transfer efficiency to the optimal transfer efficiency in the cases of the polarization transfer problems (4.8) for $\xi = 1$ (left), and (4.9) for $\xi_a = 1$ and $\xi_c = 0.75$ (right). The error in these examples is the difference between the pseudospectral transfer efficiency and the analytic optimal efficiency.

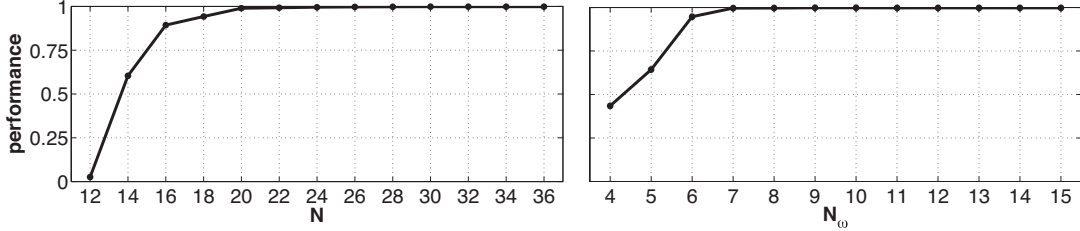


Figure 5.2: The numerical convergence of the Bloch system is shown in both the time and parameter domains. The performance of the optimal broadband $\pi/2$ pulse converges to unity as the number of discretizations, N , or samples, N_ω , independently get large. (Parameter values: maximum rf amplitude = 20 kHz; bandwidth = $[-20, 20]$ kHz; duration = $100\mu s$; fixed $N_\omega = 8$, left; fixed $N = 30$, right).

of the pseudospectral method for optimal control, when implemented for the problems of polarization transfer given by (4.8) and (4.9). The error, when compared to the analytic ROPE and CROP solutions, respectively, the error quickly decreases to zero. Figure 5.2 shows the rapid convergence of the multidimensional pseudospectral method in both the discretization (time) and sampling (parameter) dimensions for a broadband $\pi/2$ pulse maximizing the terminal x value across the ensemble, see (4.4) for $\alpha_1 = 1$, $\alpha_2 = 0$. As the order of discretization (N) and/or sampling (N_s) increase, the method yields an objective ($\varphi(T) = M_x(T, \omega, 1)$) that converges to the maximum value of unity. The low order of approximation is a characteristic of the orthogonal approximations at the heart of the numerical method. Although such empirical figures are convincing, we now show this convergence in a more rigorous fashion.

5.2 Preliminaries

The results in this section will provide the foundation on which we can analyze the feasibility, convergence, and consistency of the pseudospectral approximation method for optimal control problems. We begin by presenting several key established results in polynomial approximation theory and the natural vector extensions. With these inequalities, we are able then to prove feasibility and convergence. We define an optimal solution to Problem 1 as any feasible solution that achieves the optimal cost $J(x^*, u^*) = J^*$. We use this definition of an optimal solution within the subsequent preliminaries and the main result. To this end, the last lemma of this section introduces the error in the cost due to interpolation.

Remark 3. *Given Problem 1, $x \in H_n^\alpha(\Omega)$. Since $x(t)$ exists and $f \in C_n^{\alpha-1}$, all the derivatives $x^{(k)} \in C_n^0$, $\forall k = 0, 1, \dots, \alpha$ exist and are square integrable on the compact domain Ω , $x^{(k)} \in L_n^2(\Omega)$. Therefore, $x \in H_n^\alpha(\Omega)$.*

Lemma 1 (Interpolation Error Bounds [10], p. 289). *If $h \in H^\alpha(\Omega)$, the following hold with $c_1, c_2, c_3, c > 0$.*

(a) *The interpolation error is bounded,*

$$\|h - \mathcal{I}_N h\|_2 \leq c_1 N^{-\alpha} \|h\|_{(\alpha)}.$$

(b) *The error between the exact derivative and the derivative of the interpolation is bounded,*

$$\|\dot{h} - \mathcal{D}_N h\|_2 \leq c_2 N^{1-\alpha} \|h\|_{(\alpha)}.$$

The same bound holds for the discrete $L^2(\Omega)$ norm,

$$\|\dot{h} - \mathcal{D}_N h\|_N \leq c_3 N^{1-\alpha} \|h\|_{(\alpha)}.$$

(c) *The error due to quadrature integration is bounded,*

$$\left| \int_{-1}^1 h(t) dt - \sum_{k=0}^N h(t_k) w_k \right| \leq c N^{-\alpha} \|h\|_{(\alpha)},$$

where t_k is the k^{th} LGL node and w_k is the corresponding k^{th} weight for LGL quadrature.

Lemma 2. If $h \in H_n^\alpha(\Omega)$, i.e., an n -vector valued Sobolev space, $h = (h_1 \ h_2 \ \dots \ h_n)^T$, $h_i \in H^\alpha(\Omega)$, $i = 1, 2, \dots, n$.

(a) The vector-valued extension of Lemma 1a is, by the triangular inequality on the $L_n^2(\Omega)$ norm,

$$\|h - \mathcal{I}_N h\|_2 \leq \sum_{i=1}^n \|h_i - \mathcal{I}_N h_i\|_2 \leq \sum_{i=1}^n c_i N^{-\alpha} \|h_i\|_{(\alpha)}.$$

(b) Similarly, 1b can be extended,

$$\begin{aligned} \|\dot{h} - \mathcal{D}_N h\|_2 &\leq \sum_{i=1}^n \|\dot{h}_i - \mathcal{D}_N h_i\|_2 \leq \sum_{i=1}^n c_i N^{1-\alpha} \|h_i\|_{(\alpha)} \\ &\leq c N^{1-\alpha}, \end{aligned}$$

which again also holds for the discrete $L_n^2(\Omega)$ norm.

Proposition 1 (Feasibility). Given a solution (x, u) of Problem 1, then Problem 2 has a feasible solution, (\bar{x}, \bar{u}) , which are the corresponding interpolation coefficients.

Proof. Given the feasible solution (x, u) , let $(\mathcal{I}_N x, \mathcal{I}_N u)$ be the polynomial interpolation of this solution at the LGL nodes. Our aim is to show that the coefficients of this interpolation satisfy (5.7)-(5.9) of Problem 2. Consider the constraints imposed by the dynamics in (5.7). Because the discrete norm is evaluated only at the interpolation points,

$$\begin{aligned} \|f(\mathcal{I}_N x, \mathcal{I}_N u) - \mathcal{D}_N x\|_N &= \|f(x, u) - \mathcal{D}_N x\|_N \\ &= \|\dot{x} - \mathcal{D}_N x\|_N \\ &\leq c_d N^{1-\alpha} \end{aligned}$$

where the last step is given by Lemma 2b. Therefore, the interpolation coefficients (\bar{x}, \bar{u}) satisfy the dynamics of Problem 2 in (5.7). We can easily show that the path

constraints are also satisfied because $g(x(t), u(t)) \leq 0$ for all $t \in \Omega$ by (5.4). Because this holds for all $t \in \Omega$, it also holds for all LGL nodes $t_k \in \Gamma^{\text{LGL}}$, i.e.,

$$g(\bar{x}_k, \bar{u}_k) = g(x(t_k), u(t_k)) \leq 0,$$

which gives (5.9). The endpoint constraints are trivially satisfied by the definition of interpolation and the presence of interpolation nodes at both endpoints. Therefore, (\bar{x}, \bar{u}) is a feasible solution to Problem 2. \square

Proposition 2 (Convergence). *Given the sequence of solutions to Problem 2, $\{(\bar{x}, \bar{u})\}_N$, then the sequence of corresponding interpolation polynomials, $\{(\mathcal{I}_N x, \mathcal{I}_N u)\}$, has a convergent subsequence, such that*

$$\lim_{N_j \rightarrow \infty} (\mathcal{I}_{N_j} x, \mathcal{I}_{N_j} u) = (\mathcal{I}_\infty x, \mathcal{I}_\infty u),$$

which is a feasible solution to Problem 1.

Proof. Given that (\bar{x}, \bar{u}) is a feasible solution of Problem 2, it satisfies (5.7)-(5.9). Our goal is to show (i) that the sequence of solutions, $\{(\mathcal{I}_N x, \mathcal{I}_N u)\}_N$, has a convergent subsequence and (ii) that its limit is a feasible solution of Problem 1, satisfying (5.2)-(5.4). We first show (ii) by the assumption of (i).

(ii) Explicitly writing out the discrete norm in (5.7) gives

$$\left(\sum_{k=0}^N \sum_{i=1}^n (f_i(\mathcal{I}_N x, \mathcal{I}_N u) - \mathcal{D}_N x_i)^2(t_k) \right)^{1/2} \leq c_d N^{1-\alpha}.$$

Because f is continuous, it satisfies

$$\begin{aligned} \lim_{N \rightarrow \infty} (f_i(\mathcal{I}_N x, \mathcal{I}_N u) - \mathcal{D}_N x_i)(t_k) \\ = (f_i(\mathcal{I}_\infty x, \mathcal{I}_\infty u) - (\mathcal{I}_\infty x)')(t_k) = 0, \end{aligned}$$

therefore,

$$\frac{d}{dt}(\mathcal{I}_\infty x)(t_k) = f(\mathcal{I}_\infty x, \mathcal{I}_\infty u)(t_k), \quad (5.11)$$

which states that $(\mathcal{I}_\infty x, \mathcal{I}_\infty u)$ satisfies the dynamics in (5.2) at the interpolation nodes. Moreover, as $N \rightarrow \infty$, the LGL nodes $t_k \in \Gamma^{\text{LGL}}$ are dense in Ω , which further

shows that $(\mathcal{I}_\infty x, \mathcal{I}_\infty u)$ satisfies the dynamics of Problem 1 at all points on the interval Ω . Similarly, one can prove that this solution satisfies the path constraints because the LGL nodes become dense in Ω as $N \rightarrow \infty$ and $g(\bar{x}_k, \bar{u}_k) = g(x(t_k), u(t_k)) \leq 0$ at all LGL nodes. Again, the endpoint constraints are met exactly because the LGL grid has nodes at the endpoints.

(i) We now return to establish the existence of a convergent subsequence. The sequence $\{\mathcal{I}_N x\}$ is a sequence of bounded polynomials on a compact domain, therefore, for each finite N , $\mathcal{I}_N x \in H_n^\alpha(\Omega)$. In the limit, we showed above in (5.11) that $d/dt \mathcal{I}_\infty x$ matches $f(\mathcal{I}_\infty x, \mathcal{I}_\infty u)$ across the entire interval so that $\mathcal{I}_\infty x \in C_n^\alpha$. With the boundedness of the interpolating polynomials and the compactness of Ω , Rellich's Theorem (cf., e.g., [17], p. 272) gives that there is a subsequence $\{\mathcal{I}_{N_j} x\}$ which converges in $H_n^{\alpha-1}(\Omega)$. The same is true for the control interpolating polynomial. Therefore, there exists at least one limit point of the function sequence $\{(\mathcal{I}_N x, \mathcal{I}_N u)\}$ which we denote $(\mathcal{I}_\infty x, \mathcal{I}_\infty u)$. \square

Lemma 3. *Given (x, u) , where $x \in H_n^\alpha(\Omega)$ and $u \in H_m^\alpha(\Omega)$, and the corresponding interpolation coefficients, (\bar{x}, \bar{u}) , then the error in the cost functionals defined in (5.1) and (5.6) due to interpolation is given by,*

$$|J(x, u) - \bar{J}(\bar{x}, \bar{u})| \leq cN^{-\alpha}.$$

Remark 4. *Notice that (x, u) and (\bar{x}, \bar{u}) are not required to be a feasible solutions to Problem 1 and 2, respectively. This result characterizes the error due to interpolation.*

Proof. From (5.2) and (5.7) since $\varphi(x(1)) = \varphi(\bar{x}_N)$,

$$|J(x, u) - \bar{J}(\bar{x}, \bar{u})| = \left| \int_{-1}^1 \mathcal{L}(x, u) dt - \sum_{k=0}^N \mathcal{L}(\bar{x}_k, \bar{u}_k) w_k \right|.$$

Since $\mathcal{L} \in C^\alpha$, $x \in H_n^\alpha(\Omega)$, and $u \in H_m^\alpha(\Omega)$, the composite function $\tilde{\mathcal{L}}(t) = \mathcal{L}(x(t), u(t)) \in H^\alpha(\Omega)$. Let $\mathcal{L}_k = \mathcal{L}(\bar{x}_k, \bar{u}_k)$. Substituting these definitions and employing Lemma 1c, we obtain

$$\left| \int_{-1}^1 \tilde{\mathcal{L}}(t) dt - \sum_{k=0}^N \mathcal{L}_k w_k \right| \leq cN^{-\alpha} \|\tilde{\mathcal{L}}(t)\|_{(\alpha)}.$$

Because $\tilde{\mathcal{L}} \in H^\alpha(\Omega)$, $\|\tilde{\mathcal{L}}(t)\|_{(\alpha)}$ is bounded, from which the result follows. \square

5.3 Main Result

Theorem 2 (Consistency). *Suppose Problem 1 has an optimal solution (x^*, u^*) . Given a sequence of optimal solutions to Problem 2, $\{(\bar{x}^\dagger, \bar{u}^\dagger)\}_N$, then the corresponding sequence of interpolating polynomials, $\{(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger)\}_N$, has a limit point, $(\mathcal{I}_\infty x^\dagger, \mathcal{I}_\infty u^\dagger)$ which is an optimal solution to the original optimal control problem.*

Proof. We break the proof into four sections, employing the results from the previous section.

(i) By Proposition 1, since (x^*, u^*) is a solution to Problem 1, then for each choice of N , the corresponding interpolation coefficients, (\bar{x}^*, \bar{u}^*) , are a feasible solution to Problem 2. By the definition of optimality of $(\bar{x}^\dagger, \bar{u}^\dagger)$,

$$\bar{J}(\bar{x}^\dagger, \bar{u}^\dagger) \leq \bar{J}(\bar{x}^*, \bar{u}^*). \quad (5.12)$$

(ii) By Proposition 2, the limit point of the polynomial interpolation of the discrete optimal solution to Problem 2, $\lim_{N \rightarrow \infty} (\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger) = (\mathcal{I}_\infty x^\dagger, \mathcal{I}_\infty u^\dagger)$, is a feasible solution of Problem 1. Therefore, we have, by the definition of the optimality of (x^*, u^*) and the continuity of J ,

$$\begin{aligned} J(x^*, u^*) &\leq \lim_{N \rightarrow \infty} J(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger) \\ &= J(\mathcal{I}_\infty x^\dagger, \mathcal{I}_\infty u^\dagger). \end{aligned} \quad (5.13)$$

(iii) Using Lemma 3, we can bound the error in the cost between the optimal solution of Problem 1, (x^*, u^*) , and the corresponding interpolating coefficients, (\bar{x}^*, \bar{u}^*) , as

$$|J(x^*, u^*) - \bar{J}(\bar{x}^*, \bar{u}^*)| \leq c_1 N^{-\alpha}. \quad (5.14)$$

Similarly, we can bound the error in the cost between the optimal solution of Problem 2, $(\bar{x}^\dagger, \bar{u}^\dagger)$, and the polynomial interpolation of this solution, $(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger)$, as

$$|J(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger) - \bar{J}(\bar{x}^\dagger, \bar{u}^\dagger)| \leq c_2 N^{-\alpha}. \quad (5.15)$$

Recall that Lemma 3 does not require $(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger)$ to be a feasible solution of Problem 1. From (5.14) and (5.15),

$$\lim_{N \rightarrow \infty} \bar{J}(\bar{x}^*, \bar{u}^*) = J(x^*, u^*), \quad (5.16)$$

$$\lim_{N \rightarrow \infty} [J(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger) - \bar{J}(\bar{x}^\dagger, \bar{u}^\dagger)] = 0. \quad (5.17)$$

(iv) We are now ready to assemble the various pieces of this proof. Combining (5.16) and (5.12) we have,

$$\lim_{N \rightarrow \infty} \bar{J}(\bar{x}^\dagger, \bar{u}^\dagger) \leq \lim_{N \rightarrow \infty} \bar{J}(\bar{x}^*, \bar{u}^*) = J(x^*, u^*).$$

Adding the result from (5.13),

$$\lim_{N \rightarrow \infty} \bar{J}(\bar{x}^\dagger, \bar{u}^\dagger) \leq J(x^*, u^*) \leq \lim_{N \rightarrow \infty} J(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger). \quad (5.18)$$

Since the difference between the left and right sides, as given by (5.17), decreases to zero as $N \rightarrow \infty$, the quantities $\bar{J}(\bar{x}^\dagger, \bar{u}^\dagger)$ and $J(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger)$ converge to $J(x^*, u^*)$, i.e.,

$$\begin{aligned} 0 &\leq \lim_{N \rightarrow \infty} [J(x^*, u^*) - \bar{J}(\bar{x}^\dagger, \bar{u}^\dagger)] \\ &\leq \lim_{N \rightarrow \infty} [J(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger) - \bar{J}(\bar{x}^\dagger, \bar{u}^\dagger)] = 0. \end{aligned}$$

Thus the optimal discrete cost $\bar{J}(\bar{x}^\dagger, \bar{u}^\dagger)$ of Problem 2 and the continuous cost $J(\mathcal{I}_N x^\dagger, \mathcal{I}_N u^\dagger)$ of the corresponding interpolation polynomials converge to the optimal cost $J(x^*, u^*)$ of Problem 1. Moreover, $(\mathcal{I}_\infty x^\dagger, \mathcal{I}_\infty u^\dagger)$ is a feasible solution to Problem 1 and achieves the optimal cost. Therefore, $(\mathcal{I}_\infty x^\dagger, \mathcal{I}_\infty u^\dagger)$ is an optimal solution to Problem 1. \square

5.4 Ensemble Extension

Ensemble Control pertains to the study of a continuum of dynamical systems of the form [39],

$$\frac{d}{dt}x(t, s) = F(t, s, x(t, s), u(t)), \quad (5.19)$$

which is indexed by a parameter vector that exhibits variation within an interval, $s \in \mathcal{S} \subset \mathbb{R}^d$ but controlled by the open loop input $u(t)$. Such systems arise from environmental interactions, uncertainty, or inherent variability that induces inhomogeneity in the characteristic parameters of the dynamics. An optimal ensemble control problem is formulated by replacing the dynamics with the ensemble dynamics in (5.19) and the cost with,

$$J = \left(\int_{\mathcal{S}} \varphi(x(1, s)) + \int_{-1}^1 \mathcal{L}(x(t, s), u(t)) dt \right) ds, \quad (5.20)$$

and the end and path constraints are extended in a straightforward manner. The method employs $d + 1$ dimensional interpolating polynomials to represent x and u with the approximate dynamics (compare to (3.7)) given by [56],

$$\frac{d}{dt} I_{N \times N_{s_1} \times \dots \times N_{s_d}} x(t, \mathbf{s}_j) = \sum_{k=0}^N D_{ik} \bar{x}_{kj_1 \dots j_d}, \quad (5.21)$$

where $\mathbf{s} = (s_1, s_2, \dots, s_d)' \in \mathcal{S} \subset \mathbb{R}^d$. This extension hinges upon the lack of time dependence in the new dimensions of the problem (d parameter dimensions). Propositions 1 and 2 can then be extended in a straightforward manner by incorporating additional dynamics constraints that act in parallel. Lemma 3 will include gaussian quadrature approximations of both the s and t integrals. With these limited modifications, the approach above guarantees the convergence of the multidimensional pseudospectral method applied to optimal ensemble control problems.

Chapter 6

Conclusions

This thesis has presented a computational framework for optimal control of inhomogeneous ensembles. A framework implies an approach that includes interconnected components of theory, computation, and application. In this work the innovations in computation are mutually motivated by theory and application creating a systematic and mathematically directed implementation of control and systems theory to applications in the natural sciences and engineering.

We began by motivating the study of inhomogeneous ensembles through applications arising in - but not limited to - the control of quantum, biological, and uncertain systems. These large-scale complex systems can be best modeled as a family of dynamical systems, which are structurally similar, but exhibit variation in characteristic parameters. In these applications direct control of each sub-member is not possible and only a single global control is available to steer these systems along a desired state evolution. Moreover, it is either expensive or infeasible to observe the full-state for feedback control. Practical problems for these applications yield a novel optimal ensemble control problem to design open-loop controls to minimize an adapted Bolza cost functional subject to the dispersive dynamics of the ensemble system. Solving such optimal ensemble control problems is, in general, analytically intractable, which requires us to develop a robust computational approach.

In the review of ensemble control literature we indicated that the controllability analysis of ensemble Bloch systems was based on the concept of polynomial approximation. This dependence on polynomial approximation is then reflected in the pseudospectral computational method chosen, adapted, and developed to solve optimal ensemble control problems. This method employs both orthogonal and interpolating polynomial

expansions to represent and discretize the states and controls of the continuous-time optimal control problem. The method approximates the original problem with a nonlinear programming problem, which can be solved with any choice of nonlinear optimization solver.

We then presented several results from the control of quantum systems that highlight the ease and effectiveness of the computational framework. In problems of broadband excitation and inversion, modeled by the inhomogeneous Bloch equations, we designed several high performing controls, or pulses, that were faithfully reproduced in experiments for NMR in liquids. When a single inversion pulse was incorporated in a real protein experiment (HSQC 2D spectra), the sensitivity enhancement was 20 times higher than conventional methods. In relaxation optimized quantum systems we were able to recover very closely the analytically derived optimal efficiency of polarization transfer.

We then outlined a novel proof for the convergence of the pseudospectral method for general systems. A convergence proof and rates had, prior to our work, only been available for certain classes of systems, which did not include the bilinear form of quantum control problems. We generalize the approach for this analysis by showing that the sequence of solutions to the nonlinear programming problem created by the multidimensional pseudospectral approximations, as the order of approximation (N) increases, converges to the optimal solution of the original optimal control problem.

6.1 Future Work

Ensemble control is a very new area of research and promises to deliver advancements in systems and control theory as well as to a wide variety of applications across science and engineering.

Ensemble controllability has been studied for bilinear systems, as arising from the control of quantum systems, and time-varying linear systems. The approaches taken for these two classes of systems are substantially different and suggests that the further analysis for general nonlinear systems will require a new techniques and fundamental

theory. In addition the notion of approximate controllability and reachable sets are of particular interest in the case where full state-space controllability is unavailable.

There are also many avenues of future research to explore in the area of computational methods for ensemble control. Even within the class of spectral and pseudospectral methods, the choice of orthogonal function family and interpolation nodes can make a difference in the effectiveness of a method. There are opportunities to develop approaches based on new geometries, such as spherical spectral methods for optimal ensemble control.

One of the most exciting frontiers of ensemble control is the wide array of new applications. In biological systems, we find ensembles of interconnected entities, which inspires new theoretical analysis of connected inhomogeneous ensembles. The control of uncertain systems with ensemble control can contribute to, for example, flight control and robotic manufacturing. As we push the limits of our understanding of large-scale nano- and micro-leveled systems, ensemble control will continue to find new areas of application and the practical constraints of these systems will require a strong computational framework based on optimal control of inhomogeneous ensembles.

Appendix A

Pseudospectral Method Supplement

A.1 Orthogonal Polynomials

Given a non-negative weight function $w(t) \geq 0$, $\int_a^b w(t)dt > 0$, and a weighted inner product $f, g \in L_w^2(a, b)$,

$$\langle f, g \rangle_w = \int_a^b f(t)g(t)w(t)dt,$$

it is possible to create an orthogonal basis, $\{\phi_k\}$, using the Gram-Schmidt process, i.e.,

$$\langle \phi_i, \phi_j \rangle_w \propto \delta_{ij}.$$

Furthermore, orthogonalizing the non-negative powers of t yields a set of orthogonal *polynomials*,

$$(1, t, t^2, \dots, t^N; w(t)) \Rightarrow \{p_k\}$$

where $p_k \in \mathbb{P}_{\mathbb{N}}$. Legendre polynomials, $\{L_k(t)\}$, are derived with unit weight function, $w(t) = 1$. Therefore, the Gram-Schmidt process for the Legendre polynomials is given by $L_0(t) = 1$ and

$$L_k(t) = t^k - \sum_{i=0}^{k-1} \underbrace{\frac{\langle t^k, L_i(t) \rangle}{\langle L_i(t), L_i(t) \rangle}}_{\text{project } t^k \text{ onto } L_i(t)} L_i(t) = t^k - \sum_{i=0}^{k-1} \frac{\int_{-1}^1 t^k L_i(t) dt}{\int_{-1}^1 L_i^2(t) dt} L_i(t)$$

for $k > 0$. A different orthogonal polynomial family would use a different weight, w , but the process would be similar.

A.2 Legendre Polynomial Properties for Optimal Control

Recall that the ability of spectral methods to convert a differential equation into an algebraic equation is the feature which makes them powerful tools for problems such as those of optimal control. Legendre polynomials, $L_k(t)$, obey a recursion relation,

$$L_{k+1}(t) = \frac{2k+1}{k+1}tL_k(t) - \frac{k}{k+1}L_{k-1}(t) \quad (\text{A.1})$$

and also the differential relation,

$$[(1-t^2)L'_k(t)]' + k(k+1)L_k(t) = 0 \quad (\text{A.2})$$

These two relations illustrate how for function x expanded in terms of L_k , we can express $x'(t) = \dot{x}$ in terms of L_k as well - rather than in terms of L'_k . Hence, $\dot{x} = f(x)$ is now an algebraic equation since both sides can be written as an expansion using L_k as basis functions. Other useful properties of Legendre polynomials are

$$L_k(\pm 1) = (\pm 1)^k \quad (\text{A.3})$$

$$L'_k(\pm 1) = \frac{(\pm 1)^{k+1}k(k+1)}{2} \quad (\text{A.4})$$

A.3 Lagrange Interpolating Polynomials

Any interpolating polynomial can be represented by the Lagrange polynomial basis. The k^{th} Lagrange polynomial is characterized by taking unit value at the k^{th} interpolation node and zero at all other nodes, which is effectively a shifted Kronecker delta function, i.e. $\ell_k(t_i) = \delta_{ki}$. The Lagrange polynomials can be written in several ways, but the most transparent is the following fractional product of the interpolation nodes,

$$\ell_k(t) = \frac{(t - t_0) \cdots (t - t_{k-1})(t - t_{k+1}) \cdots (t - t_N)}{(t_k - t_0) \cdots (t_k - t_{k-1})(t_k - t_{k+1}) \cdots (t_k - t_N)} \quad (\text{A.5})$$

$$= \prod_{\substack{i=0 \\ i \neq k}}^N \frac{(t - t_i)}{(t_k - t_i)} . \quad (\text{A.6})$$

A.4 Optimal Interpolation Nodes

The optimality of a specific choice of interpolation nodes can be quantified by

$$\|x - I_N x\|_\infty \leq (1 + \Lambda_N(\Gamma)) \|x - p_N^*(x)\|_\infty , \quad (\text{A.7})$$

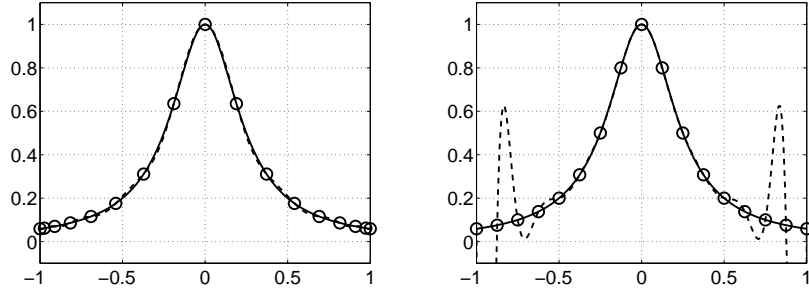


Figure A.1: Interpolation approximations of the function $f(t) = 1/(16t^2 + 1)$ using an LGL and uniform grid, respectively, with $N = 16$.

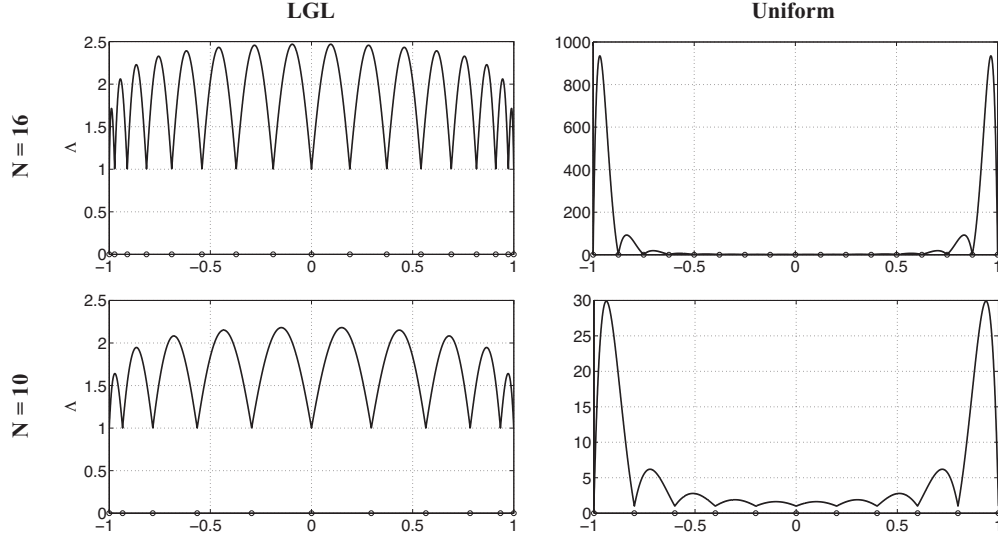


Figure A.2: The dramatic difference in interpolation error apparent in the Lebesgue constant for interpolations based on LGL and uniform grids.

where $p_N^*(x)$ is the best approximating polynomial with respect to the uniform norm and $\Lambda_N(\Gamma)$ is the Lebesgue constant defined by

$$\Lambda_N(\Gamma) = \max_{t \in [-1, 1]} \sum_{k=0}^N |\ell_k(t)|, \quad (\text{A.8})$$

with $\ell_k(t)$ the k^{th} Lagrange polynomial for the interpolation grid Γ . The Lebesgue constant, then, gives the maximum cumulative excursion from zero of the Lagrange polynomial family along the time axis. An example for the characteristically difficult function

$$f(t) = \frac{1}{16t^2 + 1},$$

is shown in terms of the interpolation (Figure A.1) and Lebesgue constant (Figure A.2).

Although a closed form for the Lebesgue constant is not in the literature, as $N \rightarrow \infty$ the Chebyshev-Gauss grid (close to the LGL nodes) yields,

$$\Lambda_N(\Gamma_{\text{CG}}) = \frac{2}{\pi} \log N + \frac{2}{\pi} \left(\gamma + \log \frac{8}{\pi} - \frac{2}{3} \right) + O\left(\frac{1}{\log N}\right)$$

which in this limit is asymptotic to the Lebesgue constant of the optimal interpolation grid [62],

$$\Lambda_N(\Gamma_{\text{CG}}) = \frac{2}{\pi} \log N + \frac{2}{\pi} \left(\gamma + \log \frac{4}{\pi} \right) + O\left(\left(\frac{\log \log N}{\log N} \right)^2 \right).$$

A.5 Lagrange Polynomial written in terms of Legendre Polynomial

Define $w(t) = \prod_{i=0}^N (t - t_i)$. Taking the derivative,

$$w'(t) = \sum_{k=0}^N \prod_{\substack{i=0 \\ i \neq k}}^N (t - t_i) \quad \Rightarrow \quad w'(t_k) = \prod_{\substack{i=0 \\ i \neq k}}^N (t_k - t_i) \quad (\text{A.9})$$

We can now express (A.5) from Appendix A.3 as,

$$\ell_k(t) = \frac{w(t)}{(t - t_k)w'(t_k)} \quad (\text{A.10})$$

Recall that the LG nodes (LGL nodes excluding the endpoints) $\{t_1, \dots, t_{N-1}\}$ are zeros of $L'_N(t)$, therefore $L'_N(t) = (t - t_1) \dots (t - t_{N-1})$. We can then write $w(t)$ in terms of the N degree Legendre polynomial.

$$\begin{aligned} w(t) &= (t - t_0) \underbrace{(t - t_1) \dots (t - t_{N-1})}_{L'_N(t)} (t - t_N) \\ &= (t^2 - 1)L'_N(t) \end{aligned} \quad (\text{A.11})$$

Combining (A.11) with the Legendre derivative relation (A.2) from Appendix A.2,

$$w'(t_k) = [(t_k^2 - 1)L'_N(t_k)]' = N(N + 1)L_N(t_k) \quad (\text{A.12})$$

Substituting (A.11) and (A.12) into (A.10) we yield an expression for the Lagrange interpolating functions in terms of the Legendre polynomials.

$$\ell_k(t) = \frac{1}{N(N+1)L_N(t_k)} \frac{(t^2-1)L'_N(t)}{t-t_k}$$

Once we have the Lagrange polynomials in terms of the Legendre polynomials we can analytically compute the weights for LGL quadrature integration,

$$w_k = \int_{-1}^1 \ell_k(t) dt = \frac{2}{N(N+1)} \frac{1}{[L_N(t_k)]^2}, \quad i = 0, 1, \dots, N. \quad (\text{A.13})$$

A.6 Derivative Matrix

$D_{ik} = \dot{\ell}_k(t_i)$ is an $(N+1) \times (N+1)$ matrix. Taking the time derivative of (3.6), and using the Legendre relation (A.2) from Appendix A.2 for the derivative of the numerator yields,

$$\left[\frac{\partial}{\partial t} \ell_k(t) \right]_{t=t_i} = \frac{1}{N(N+1)L_N(t_k)} \left[\frac{N(N+1)L_N(t_i)}{t_i - t_k} - \frac{(t_i^2 - 1)L'_N(t_i)}{(t_i - t_k)^2} \right] \quad (\text{A.14})$$

For any $i \neq k$, the second term in the brackets is zero, since $t = t_i$ is a zero of $(t^2 - 1)L'_N(t)$. Canceling terms, yields the first component of the derivative matrix in (3.8). For $i = k$, we utilize l'Hopital's rule for each term (we use the rule twice for the second term).

$$\lim_{t \rightarrow t_k} \frac{N(N+1)L_N(t)}{(t - t_k)} = \lim_{t \rightarrow t_k} \frac{N(N+1)L'_N(t)}{1} = N(N+1)L'_N(t_k) \quad (\text{A.15})$$

$$\lim_{t \rightarrow t_k} \frac{(t^2 - 1)L'_N(t)}{(t - t_k)^2} = \lim_{t \rightarrow t_k} \frac{N(N+1)L_N(t)}{2(t - t_k)} = \lim_{t \rightarrow t_k} \frac{N(N+1)L'_N(t)}{2} = \frac{N(N+1)L'_N(t_k)}{2} \quad (\text{A.16})$$

For $i = k \neq 0, N$, $L'_N(t_k) = 0$ which indicates that $\dot{\ell}_k(t_k) = 0$ if $k \neq 0, N$. Substituting the values of $L_N(\pm 1)$ and $L'_N(\pm 1)$ given by equations (A.3-A.4) in Appendix A.2 and combining these two terms,

$$\frac{\partial \ell_k(t_k)}{\partial t} = (\pm 1) \frac{N(N+1)}{4} \quad k \in \{0, N\} \quad (\text{A.17})$$

Therefore, the elements of D are as given in (3.8),

$$D_{ik} = \begin{cases} \frac{L_N(t_i)}{L_N(t_k)} \frac{1}{t_i - t_k} & i \neq k \\ -\frac{N(N+1)}{4} & i = k = 0 \\ \frac{N(N+1)}{4} & i = k = N \\ 0 & \text{otherwise.} \end{cases}$$

The matrix D is the first differentiation matrix. Optimal control requires only the first derivative, as the differential equation is of order 1. Similar expressions can be computed for second, third, etc. differentiation matrices.

Appendix B

AMPL Examples

B.1 Single Spin Bloch Optimization

Problem Definition: $\max x(T)$, $T = \pi/2$, subject to $u(t)^2 + v(t)^2 \leq 1, t \in [0, T]$ and

$$\frac{d}{dt} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 & 0 & u \\ 0 & 0 & -v \\ -u & v & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}, \quad \begin{bmatrix} x(0) \\ y(0) \\ z(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \quad (\text{B.1})$$

```
param N > 0 integer;
```

```
param A > 0;
```

```
param T > 0;
```

```
param x0;
```

```
param y0;
```

```
param z0;
```

```
set nodes := 1..(N+1);
```

```
param D {nodes,nodes};
```

```
var x {nodes} >= -1, <= 1;
```

```
var y {nodes} >= -1, <= 1;
```

```
var z {nodes} >= -1, <= 1;
```

```

var u {nodes} >=-A, <=A;
var v {nodes} >=-A, <=A;

maximize cost: x[N+1];

subject to dynamics_x {t in nodes}:
    u[t]*z[t] = (2/T)*(sum{k in nodes} D[t,k]*x[k]);

subject to dynamics_y {t in nodes}:
    -v[t]*z[t] = (2/T)*(sum{k in nodes} D[t,k]*y[k]);

subject to dynamics_z {t in nodes}:
    -u[t]*x[t] + v[t]*y[t] = (2/T)*(sum{k in nodes} D[t,k]*z[k]);

subject to initialConditions_x: x[1] = x0;
subject to initialConditions_y: y[1] = y0;
subject to initialConditions_z: z[1] = z0;

subject to amplitudeBound {t in nodes}: u[t]^2+v[t]^2 <= A^2;

data;
param N := 10;
param A := 1;
param T := pi/2;

param x0 := 0;
param y0 := 0;
param z0 := 1;

param D :
    1    2    3    4    5    6    7    8    9    10    11 :=
    1    -27.50    37.20   -14.88    8.49   -5.64    4.06   -3.06    ...
    2     -6.17     0.00     8.73   -4.07     2.53   -1.77     1.31    ...
    3      1.44   -5.11     0.00    5.25   -2.53     1.61   -1.14    ...

```

4	...							
5	...							
6	...							
7	...							
8	...							
9	...							
10	...							
11	0.50	-1.26	1.79	-2.35	3.06	-4.06	5.64	...

;

B.2 Broadband Spin Bloch Optimization

Problem Definition: $\max \int_{\mathcal{S}} x(T, \omega) d\omega$, $0 \leq T \leq 2\pi$, $\mathcal{S} = [-1, 1]$ subject to $u(t)^2 + v(t)^2 \leq 1, t \in [0, T]$,

$$\frac{d}{dt} \begin{bmatrix} x(t, \omega) \\ y(t, \omega) \\ z(t, \omega) \end{bmatrix} = \begin{bmatrix} 0 & -\omega & u \\ \omega & 0 & -v \\ -u & v & 0 \end{bmatrix} \begin{bmatrix} x(t, \omega) \\ y(t, \omega) \\ z(t, \omega) \end{bmatrix}, \quad \begin{bmatrix} x(0, \omega) \\ y(0, \omega) \\ z(0, \omega) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \quad (\text{B.2})$$

```

param N > 0 integer;
param Nw > 0 integer;
param A > 0;
param B > 0;
param Tmax > 0;

param x0;
param y0;
param z0;

set states := 1..3;
set nodes := 1..(N+1);
set dispersion := 1..(Nw+1)

```

```

param D {nodes,nodes};
param w {dispersion};
param wwts {dispersion};

var T >= 0 <= Tmax;
var M {states,dispersion,nodes} >= -1, <= 1;

var u {nodes} >=-A, <=A;
var v {nodes} >=-A, <=A;

maximize cost: ((2*B)/2)*(sum{i in dispersion} M[1,i,N+1]*wwts[i]);

subject to dynamics_x {i in dispersion, t in nodes}:
    -w[i]*M[2,i,t] + u[t]*M[3,i,t]
        = (2/T)*(sum{k in nodes} D[t,k]*M[1,i,k]);

subject to dynamics_y {i in dispersion, t in nodes}:
    w[i]*M[1,i,t] - v[t]*M[3,i,t]
        = (2/T)*(sum{k in nodes} D[t,k]*M[2,i,k]);

subject to dynamics_z {i in dispersion, t in nodes}:
    -u[t]*M[1,i,t] + v[t]*M[2,i,t]
        = (2/T)*(sum{k in nodes} D[t,k]*M[3,i,k]);

subject to initialConditions_x {i in dispersion}:
    M[1,i,1] = x0;
subject to initialConditions_y {i in dispersion}:
    M[2,i,1] = y0;
subject to initialConditions_z {i in dispersion}:
    M[3,i,1] = z0;

subject to amplitudeBound {t in nodes}: u[t]^2+v[t]^2 <= A^2;

data;

```

```
param N := 10;
param Nw := 4;
param A := 1;
param B := 1;
param Tmax := 2*pi;
```

```
param x0 := 0;
param y0 := 0;
param z0 := 1;
```

```
param w :=
  1  -1
  2  -0.5
  3  0
  4  0.5
  5  1
;
```

```
param wwts :=
  1  0.1
  2  0.54
  3  0.71
  4  0.54
  5  0.1
;
```

```
param D :
  ...
;
```

Appendix C

Bloch Equations

The Bloch equations without relaxation, $\dot{\mathcal{M}} = \mathcal{M} \times \gamma \mathcal{B}_{\text{eff}}$, utilizes the classical description of interacting electromagnetic forces, where \mathcal{M} is the spin magnetization vector, γ is the gyromagnetic ratio, the effective externally applied field is $\mathcal{B}_{\text{eff}} = (B_1 \cos(\omega_0 t + \phi), B_1 \sin(\omega_0 t + \phi), B_0)'$, $B_1(t)$ and B_0 are the amplitudes of the applied fields in the transverse plane and z direction respectively, and $\phi(t)$ is the phase angle [15]. Conventionally, the fields are given as frequencies $\gamma \mathcal{B}_{\text{eff}} = (\omega_{1x}, \omega_{1y}, \omega_0)$ and measured in units of Hertz. Using the generators of rotation,

$$\Omega_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad \Omega_y = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \quad \Omega_z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

the Bloch equations are be given by

$$\frac{d}{dt} \mathcal{M}(t) = \left[\omega_0 \Omega_z + \omega_{1y}(t) \Omega_y + \omega_{1x}(t) \Omega_x \right] \mathcal{M}(t). \quad (\text{C.1})$$

If we consider variation in the applied electromagnetic fields B_0 and B_1 , we can express (C.1) in matrix form,

$$\begin{aligned}
& \frac{d}{dt} \begin{bmatrix} \mathcal{M}_x(t, \omega, \epsilon) \\ \mathcal{M}_y(t, \omega, \epsilon) \\ \mathcal{M}_z(t, \omega, \epsilon) \end{bmatrix} \\
&= \begin{bmatrix} 0 & -(\omega_0 + \omega) & \epsilon\omega_1 \sin(\omega_0 t + \phi) \\ \omega_0 + \omega & 0 & -\epsilon\omega_1 \cos(\omega_0 t + \phi) \\ -\epsilon\omega_1 \sin(\omega_0 t + \phi) & \epsilon\omega_1 \cos(\omega_0 t + \phi) & 0 \end{bmatrix} \begin{bmatrix} \mathcal{M}_x(t, \omega, \epsilon) \\ \mathcal{M}_y(t, \omega, \epsilon) \\ \mathcal{M}_z(t, \omega, \epsilon) \end{bmatrix}
\end{aligned}$$

where $\omega \in [-\beta, \beta]$ and $\epsilon \in [1 - \delta, 1 + \delta]$, $0 \leq \delta \leq 1$. For calculation and computation, it is useful to transform the Bloch equations into the so-called rotating frame and normalize the system by a nominal pulse amplitude A_0 to yield a dimensionless equation. Solutions based on the dimensionless equation can then be scaled for a specific choice of nominal amplitude. Consider a transformation $M = \exp(-\omega_0 \Omega_z t) \mathcal{M}$. In addition we scale time with $\tau = A_0 t$. It is straightforward to show that the new state equation is given by,

$$\frac{d}{d\tau} M(\tau, \omega, \epsilon) = \left[\omega \Omega_z + \epsilon u(\tau) \Omega_y + \epsilon v(\tau) \Omega_x \right] M(\tau, \omega, \epsilon),$$

where $\tau \in [0, AT \times 2\pi]$, $\omega \in [-B, B]$, $B = \beta/A_0$, and

$$u(\tau) = \frac{\gamma B_1(\tau/A_0)}{A_0} \cos(\phi(\tau/A_0)) \quad v(\tau) = \frac{\gamma B_1(\tau/A_0)}{A_0} \sin(\phi(\tau/A_0)),$$

(all dimensionless). Note the 2π factor in the time scaling is introduced to convert from units of Hertz to radians/second. Designing the time-varying controls $u(\tau)$ and $v(\tau)$ is equivalent to the original design of amplitude $B_1(t)$ and phase $\phi(t)$.

C.1 Conversion

To be absolutely clear regarding the normalization of the Bloch equation and how to easily convert between physically meaningful parameter values and those used for computation, below is a conversion guide. In most cases the nominal amplitude A_0 is chosen as the maximum amplitude.

→ Given *real* values corresponding to maximum amplitude, A (Hz); half bandwidth, B (Hz); and duration T (s),

$$\tilde{A} = A/A_0$$

$$\tilde{B} = B/A_0$$

$$\tilde{T} = T * A_0 \times 2\pi$$

← Given *normalized* dimensionless values corresponding to maximum amplitude, \tilde{A} ; half bandwidth, \tilde{B} ; and duration \tilde{T} ,

$$A = \tilde{A} * A_0$$

$$B = \tilde{B} * A_0$$

$$T = \frac{\tilde{T}}{2\pi A_0}$$

References

- [1] A. Becker, T. Bretl, “Motion planning under bounded uncertainty using ensemble control,” *Robotics: Science and Systems (RSS)*, 2010.
- [2] M. A. Bernstein, K. F. King, X. Zhou, *Handbook of MRI Pulse Sequences*, Elsevier, Academic Press, 2004.
- [3] F. Bloch, “Nuclear Induction,” *Physical Review*, vol. 70, pp. 460-474, 1946.
- [4] J. Boyd, *Chebyshev and Fourier Spectral Methods*, Dover Ed. 2, New York, 2000.
- [5] R. Brockett, “Control of Stochastic Ensembles,” *Astrom Symposium on Control*, (B. Wittenmark, A. Rantzer, eds) Studentlitteratur, Lund Sweden, pp. 199-216, 1999.
- [6] R. Brockett, *Finite Dimensional Linear Systems*, Wiley, New York, 1970.
- [7] R. Brockett, “System theory on group manifolds and coset spaces,” *SIAM Journal on Control and Optimization*, vol. 10, no. 2, pp. 265-284, 1972.
- [8] R. Brockett, N. Khaneja, “On the stochastic control of quantum ensembles,” in *System Theory: Modeling, Analysis and Control*. Kluwer Academic Publishers, New York, 1999.
- [9] A. E. Bryson, Jr., Y.-C. Ho, *Applied Optimal Control*, Blaisdell, Waltham, 1969.
- [10] C. Canuto, M. Y. Hussaini, A. Quarteroni, T. A. Zang, *Spectral Methods*, Springer, Berlin, 2006.
- [11] H. Y. Carr, E. M. Purcell, “Effects of diffusion on free precession in nuclear magnetic resonance experiments,” *Physical Review*, vol. 94, pp. 630-638, 1954.
- [12] J. Cavanagh, W. J. Fairbrother, A. G. Palmer, N. J. Skelton, *Protein NMR Spectroscopy*, Academic Press, San Diego, 1996.
- [13] S. Conolly, D. Nishimura, A. Macovski, “Optimal control solutions to the magnetic resonance selective excitation problem,” *IEEE Transactions on Medical Imaging*, vol. MI-5, 1986.
- [14] G. Elnagar, M. A. Kazemi, M. Razzaghi, “The pseudospectral Legendre method for discretizing optimal control problems,” *IEEE Transactions on Automatic Control*, vol. 40, pp. 1793, 1995.

- [15] R. Ernst, G. Bodenhausen, A. Wokaun, *Principles of Nuclear Magnetic Resonance in One and Two Dimensions*, Oxford, Clarendon Press, 1987.
- [16] F. Fahroo, I. Ross, "Costate estimation by a Legendre pseudospectral method," *Journal of Guidance, Control, and Dynamics*, vol. 24, pp. 270-277, 2001.
- [17] G. B. Folland, *Real Analysis: Modern Techniques and Their Applications*, John Wiley & Sons, New York, 1984.
- [18] B. Fornberg, *A Practical Guide to Pseudospectral Methods*, Cambridge University Press, New York, 1998.
- [19] D. P. Frueh, et al, "Sensitivity enhancement in NMR of macromolecules by application of optimal control theory," *Journal of Biomolecular NMR*, vol. 32, pp. 23-30, 2005.
- [20] M. Goldman, *Quantum Description of High-Resolution NMR in Liquids*, Clarendon Press, Oxford, 1988.
- [21] Q. Gong, W. Kang, M. Ross, "A pseudospectral method for the optimal control of constrained feedback linearizable systems", *IEEE Transactions on Automatic Control*, vol. 51, no. 7, 1115-1129, 2006.
- [22] Q. Gong, M. Ross, W. Kang, F. Fahroo, "Connections between the covector mapping theorem and convergence of pseudospectral methods for optimal control", *Computational Optimization and Applications*, vol. 41, pp. 307-335, 2008.
- [23] U. Haeberlen, *High Resolution NMR in Solids*, Academic Press, New York, 1976.
- [24] E. L. Hahn, *Concepts in Magnetic Resonance*, Vol. 9 (2), pp. 6981, 1997.
- [25] E. L. Hahn, "Spin Echoes," *Physical Review*, vol. 80, pp. 580-594, 1950.
- [26] A. Isidori, *Nonlinear Control Systems*, London, Springer, 1995.
- [27] V. Jurdjevic, H. J. Sussmann, "Control systems on lie groups," *Journal of Differential Equations*, vol. 12, pp. 313-329, 1972.
- [28] C. T. Kehlet et al, "Improving solid-state NMR dipolar recoupling by optimal control," *Journal of the American Chemical Society*, vol. 126, pp. 10202-10203, 2004.
- [29] N. Khaneja, J.-S. Li, C. Kehlet, B. Luy, S. J. Glaser, "Broadband relaxation-optimized polarization transfer in magnetic resonance," *Proceedings of the National Academy of Sciences*, vol. 101, pp. 14742-14747, 2004.

- [30] N. Khaneja, B. Luy, and S. J. Glaser, “Boundary of quantum evolution under decoherence,” *Proceedings of the National Academy of Sciences*, vol. 100, pp. 13162, 2003.
- [31] N. Khaneja, T. Reiss, C. T. Kehlet, T. Schulte-Herbrüggen, S. J. Glaser, “Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms,” *Journal of Magnetic Resonance*, vol. 172, pp. 296-305, 2005.
- [32] N. Khaneja, T. Reiss, B. Luy, S. J. Glaser, “Optimal control of spin dynamics in the presence of relaxation,” *Journal of Magnetic Resonance*, vol. 162, pp. 311, 2003.
- [33] K. Kobzar, T. E. Skinner, N. Khaneja, S. J. Glaser, B. Luy, “Exploring the limits of broadband excitation and inversion pulses,” *Journal of Magnetic Resonance*, vol. 170, pp. 236-243, 2004.
- [34] K. Kobzar, T. E. Skinner, N. Khaneja, S. J. Glaser, B. Luy, “Exploring the limits of broadband excitation and inversion: II. Rf-power optimized pulses,” *Journal of Magnetic Resonance*, vol. 194, pp. 58-66, 2008.
- [35] M. H. Levitt, “Composite pulses,” *Progress in Nuclear Magnetic Resonance Spectroscopy*, vol. 18, pp. 61-122, 1986.
- [36] J.-S. Li, “Control of Inhomogeneous Ensembles,” PhD Thesis, Harvard University, 2006.
- [37] J.-S. Li, “Ensemble control of finite-dimensional time-varying linear systems,” *IEEE Transactions on Automatic Control*, vol. 56, No. 2, pp. 345-357, 2011.
- [38] J.-S. Li, N. Khaneja, “Control of inhomogeneous quantum ensembles,” *Physical Review A*, vol. 73, pp. 030302, 2006.
- [39] J.-S. Li, N. Khaneja, “Ensemble control of Bloch equations,” *IEEE Transactions on Automatic Control*, vol. 54 (3) pp. 528-536, 2009.
- [40] J.-S. Li, J. Ruths, “Optimal sampling and design of MR pulse sequences,” *Proceedings of 2009 NSF Engineering and Innovation Conference*, June 22-25, Honolulu, HI.
- [41] J.-S. Li, J. Ruths, D. Stefanatos, “A pseudospectral method for optimal control of open quantum systems,” *Journal of Chemical Physics*, vol. 131, pp. 164110, 2009.
- [42] J.-S. Li, J. Ruths, T.-Y. Yu, H. Arthanari, G. Wagner, “Optimal pulse design in quantum control: A unified computational method,” *Proceedings of the National Academy of Sciences*, vol. 108, no. 5, pp. 1879-1884, 2011.

- [43] G. Lindblad, "On the generators of quantum dynamical semigroups," *Communications in Mathematical Physics*, vol. 48, pp. 119-130, 1976.
- [44] H. Mabuchi, N. Khaneja, "Principles and applications of control in quantum systems," *Int. J. Robust Nonlinear Control*, vol. 15, pp. 647-667, 2004.
- [45] I. I. Maximov, J. Salomon, and G. Turinici, "A smoothing monotonic convergent optimal control algorithm for nuclear magnetic resonance pulse sequence design," *Journal of Chemical Physics*, vol. 132, 084107, 2010.
- [46] I. I. Maximov, Z. Tosner, N. C. Nielsen, "Optimal control design of NMR and dynamic nuclear polarization experiments using monotonically convergent algorithms," *Journal of Chemical Physics*, vol. 128, pp. 184505, 2008.
- [47] J. Moehlis, E. Shea-Brown, H. Rabitz, "Optimal inputs for phase models of spiking neurons," *Journal of Computational and Nonlinear Dynamics*, vol. 1 pp. 358-367, 2006
- [48] G.A. Morris, R. Freeman, "Enhancement of nuclear magnetic resonance signals by polarization transfer," *Journal of the American Chemical Society* vol. 101, pp. 760762, 1979.
- [49] M. Möttöen, R. Sousa, J. Zhang, K. B. Whaley, "High fidelity one-qubit operations under random telegraph noise", *Physical Review A*, vol. 73, pp. 022332, 2006.
- [50] J. Pauly, P. Le Roux, D. Nishimura, A. Macovski, "Parameter relations for the Shinnar-Le Roux selective excitation pulse design algorithm," *IEEE Transactions on Medical Imaging*, vol. 10, pp. 53-65, 1991.
- [51] G.S. Payne, M.O. Leach, "Surface-coil polarization transfer for monitoring tissue metabolism in vivo," *Magnetic Resonance in Medicine*, vol. 43, pp. 510-516, 2000.
- [52] L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, John Wiley & Sons, New York, 1962.
- [53] D. Rosenfeld, Y. Zur, "Design of adiabatic selective pulses using optimal control theory," *Magnetic Resonance in Medicine*, vol. 36, pp. 401-409, 1996.
- [54] I. Ross, F. Fahroo, "Legendre pseudospectral approximations of optimal control problems", in *New Trends in Nonlinear Dynamics and Control*, edited by W. Kang et al. (Springer, Berlin, 2003); pp. 327-342.
- [55] R. Ruskov, A. Korotkov, "Spectrum of qubit oscillations from generalized Bloch equations," *Physical Review B*, vol. 67, pp. 075303, 2003.

- [56] J. Ruths, J.-S. Li, “A multidimensional pseudospectral method for optimal control of quantum ensembles,” *Journal of Chemical Physics*, vol. 134, pp. 044128, 2011.
- [57] J. Ruths, J.-S. Li. “Optimal ensemble control of open quantum systems with a pseudospectral method,” *49th IEEE Conference on Decision and Control*, Atlanta, 2010.
- [58] J. Ruths, D. Stefanatos, T.-Y. Yu, J.-S. Li. “A universal computational method for optimal pulse design in NMR & MRI,” *Proceedings of 2011 NSF Engineering and Innovation Conference*, Jan. 4-7, Atlanta, GA.
- [59] A. J. Shaka, R. Freeman, “Composite pulses with dual compensation,” *Journal of Magnetic Resonance*, vol. 55, pp. 487-493, 1983.
- [60] A. J. Shaka, C. J. Lee, A. Pines, “Iterative schemes for bilinear operators; applications to spin decoupling,” *Journal of Magnetic Resonance*, vol. 77, pp. 274-293, 1988.
- [61] T. E. Skinner, T. Reiss, B. Luy, N. Khaneja, S. J. Glaser, “Application of optimal control theory to the design of broadband excitation pulses for high resolution NMR,” *Journal of Magnetic Resonance*, vol. 163, pp. 8-15, 2003.
- [62] S. Smith, “Lebesgue constants in polynomial interpolation,” *Annales Mathematicae et Informaticae*, vol. 33, pp. 109-123, 2006.
- [63] D. Stefanatos, J. Ruths, J.-S. Li, “Frictionless atom cooling in harmonic traps: A time optimal approach,” *Physical Review A*, vol. 82, pp. 063422, 2010.
- [64] H. J. Sussmann, “From the Brachystochrone to the maximum principle,” *35th IEEE Conference on Decision and Control*, Kobe, Japan, 1996.
- [65] G. Szego, *Orthogonal Polynomials*, American Mathematical Society, New York, 1959.
- [66] R. Tycko, H. Cho, E. Schneider, A. Pines, “Composite pulses without phase distortion,” *Journal of Magnetic Resonance*, vol. 61, pp. 90-101, 1985.
- [67] K. M. Ward, A. H. Aletras, R. S. Balaban, “A new class of contrast agents for MRI based on proton chemical exchange dependent saturation transfer (CEST),” *Journal of Magnetic Resonance*, vol. 43, pp. 79-87, 2000.

Vita

Justin Arthur Ernest Ruths

Date of Birth May 8, 1982

Place of Birth Calgary, Alberta

Degrees B.S. Cum Laude, Physics, May 2004
M.S. Mechanical Engineering, December 2005
M.S. Electrical Engineering, December 2009
Ph.D. Systems Science and Mathematics, May 2011

Publications J.-S. Li, J. Ruths, T.-Y. Yu, H. Arthanari, G. Wagner. *Optimal pulse design in quantum control: a unified computational method*, Proceedings of the National Academy of Sciences, vol. 108, no. 5, 1879-1884, 2011.

J. Ruths, J.-S. Li. *A multidimensional pseudospectral method for optimal control of quantum ensembles*, Journal of Chemical Physics, vol. 134, 044128, 2011.

D. Stefanatos, J. Ruths, J.-S. Li. *Frictionless atom cooling in harmonic traps: a time-optimal control approach*, Physical Review A, vol. 82, 063422, 2010.

J.-S. Li, J. Ruths, D. Stefanatos. *A pseudospectral method for optimal control of open quantum systems*, Journal of Chemical Physics, vol. 131, 164110, 2009.

May 2011

Optimal Control of Ensembles, Ruths, Ph.D. 2011